

Copyright (c) 1993 - 2000 Compugen Ltd.	Gencore version 4.5				
Run on:	February 1, 2002, 14:59:10 ; Search time 24.84 Seconds				
Title:	(without alignments)				
Perfect score:	377.193 Million cell updates/sec				
Sequence:	US-09-591-632-2_COPY_1_123				
Scoring table:	1 MSDSNQGNQNNQNYQISONG.....FNNNNNLOGIAGFQFQSQG 123				
Searched:	BIOSUM62 Gapext 10.0 , Gapext 0.5				
Total number of hits satisfying chosen parameters:	219241				
Minimum DB seq length:	0				
Maximum DB seq length:	200000000				
Post-processing:	Maximum Match 0%				
Database :	Listing first 45 summaries				
PIR_6B:*					
1: pir1:*					
2: pir2:*					
3: pir3:*					
4: pir4:*					
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.					
Result No.	Score	Query Match length	DB ID	Description	RESULT 1
1	701	100.0	685	1 EFBYS2	EFBYS2
2	166	23.7	405	2 S19355	suppressor 2 protein - yeast (<i>Saccharomyces cerevisiae</i>)
3	166	23.7	741	2 S12921	N;Alternate names: Gi-to-S transition protein; protein YDR172w
4	163.5	23.3	589	2 S71954	C;Species: <i>Saccharomyces cerevisiae</i>
5	148	21.1	419	2 T04886	C;Date: 31-Dec-1991 #text change 19-Jan-2001
6	147.5	21.0	409	2 T20847	C;Accession: S00733; JT0323; S05723
7	147.5	21.0	945	2 T21998	C;Wilson, P.G.; Culbertson, M.R.
8	147	21.0	462	1 LUD07	J. Mol. Biol. 199, 559-573, 1988
9	144	20.5	232	2 T21735	A;Title: Nucleotide sequence of the SUP2(SUP35) gene of <i>Saccharomyces cerevisiae</i> .
10	143	20.4	372	2 T164223	A;Reference number: JT0323; MUID:88329727
11	142	20.3	676	2 S41022	A;Accession: S00733
12	141	20.1	404	2 S46269	A;Molecule type: DNA
13	141	20.1	800	2 I51653	A;Residues: 1-685 <WIL>
14	140.5	20.0	236	2 T18240	A;Cross-references: EMBL:Y01129; NID:94581; PIDN:CAA30155.1; PID:94582
15	139	19.8	632	2 T02627	R;Kushnirov, V.V.; Tel-Avivanyan, M.D.; Telcov, M.V.; Surguchov, A.P.; Smirnov, V.N.
16	138	19.7	111	2 T29070	A;Submitted to the EMBL Data Library, November 1994
17	137.5	19.6	284	2 T23158	A;Reference number: S49768
18	137.5	19.6	765	2 T15447	A;Molecule type: DNA
19	137	19.5	455	2 C29349	A;Residues: 1-685 <WIL>
20	136.5	19.5	461	2 T27135	A;Cross-references: EMBL:Y00829; NID:93711; PIDN:CAA68760.1; PID:93712
21	136	19.4	462	4 S33798	A;Map position: 4R
22	136	19.4	526	1 S33799	C;Superfamily: suppressor 2 protein; translation elongation factor Tu homology
23	135	19.3	128	2 S45871	C;Keywords: duplication; GTP binding; nucleotide binding; P-loop; phosphoprotein; tan
24	135	19.3	528	2 G0127	F;1-123/Domain: A <DOM1>
25	135	19.3	867	2 T27136	F;42-119/Region: 10-residue repeats
26	135	19.3	685	2 T27135	F;124-253/Domain: charged <DOM2>
27	133	19.0	3498	2 T22330	F;159-222/Region: glutamic acid/lysine-rich
28	132	18.8	882	2 H70199	F;254-685/Domain: C <DOM4>
29	131	18.7	1162	2 T18400	F;261-409/Domain: translation elongation factor Tu homology <ETU>
					F;267-274/Region: nucleotide-binding motif A (P-loop)
					F;406-409/Region: GTP-binding NXND motif

F;273/Binding site: GTP (Lys) #status predicted

R;273

Query Match 100.0%; Score 701; DB 1; Length 685;
Best Local Similarity 100.0%; Pred. No. 4.2e-54;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSDSNQGNQQNQNYQQYSONGNNQQGNNRYQGYQAYNAQAQPAGCYYQNYQGSYQGY 60
Db 1 MSDSNQGNQQNQNYQQYSONGNNQQGNNRYQGYQAYNAQAQPAGCYYQNYQGSYQGY 60

Qy 61 QQYNPDAGYQQQNPOQSGYQYQVNPOGGTQOQFNPOGGERGNYKFNFVNLLNLOGYAQGPOQ 120
Db 61 QQYNPDAGYQQQNPOQGGYQQVNPOGGTQOQFNPOGGERGNYKFNFVNLLNLOGYAQGPOQ 120

Qy 121 SQG 123
Db 121 SQG 123

RESULT 2

S19355 hypothetical protein YCL028w - yeast (*Saccharomyces cerevisiae*)
N:Alternative names: hypothetical protein YCL181
C:Species: *Saccharomyces cerevisiae*
C:Accession: S19355; 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 29-Oct-1999
R:Hollenberg, C.P.; Kleinhans, U.; Lutzenkirchen, K.; Ramezani Rad, M.; xu, G.
submitted to the Protein Sequence Database, March 1992

A:Reference number: S19350

A:Accession: S19355

A:Molecule type: DNA

A:Residues: 1-405 <HOL>

A:Cross-references: EMBL:X59720; NID:91907116; PIDN:CAA42357.1; PID:e264430; PID:95336;
R:Ramezani Rad, M.; Lutzenkirchen, K.; xu, G.; Kleinhans, U.; Hollenberg, C.P.
Year 7, 533-538, 1991

A:Title: The complete sequence of a 11,953 bp fragment from C1G on chromosome III encompassing the yeast gene for the 18S rRNA.

A:Reference number: S17471; MUID:9137317

A:Accession: S17471

A:Status: translation not shown

A:Residues: 1-186 <RAM>

A:Cross-references: EMBL:SS8084

A:Molecule type: DNA

A:Residues: 1-186 <RAM>

R:Trueheart, J.; Boeke, J.D.; Fink, G.R.
Mol. Cell. Biol. 7, 2316-2328, 1987

A:Title: Two genes required for cell fusion during yeast conjugation: evidence for a pheromone response element.

A:Reference number: A93093; MUID:8728684

A:Accession: C27061

A: Molecule type: DNA

A:Residues: 1-180, 'T', 182-322, 'VHSOLNLPWQVPTWAITPIQIRVMGA NR LMSMDRN', 'RMVNSNPMS TEDRNTAE'

A:Cross-references: EMBL:M16717; NID:9171534; PIDN:AA34615.1; PID:9171536

C:Genetics:
A:Map position: 3L

Query Match 23.7%; Score 166; DB 2; Length 405;
Best Local Similarity 34.4%; Pred. No. 1.6e-07; Mismatches 55; Conservative 7; Indels 54; Gaps 9;

Qy 2 SDSNQGNQQNQNYQQYSONGNNQQGNNRYQGYQAYNAQAQPAGCYYQNYQGSYSG----- 54
Db 219 NNQNSNSNSQGQNQSYQNGNQNS----QGYN----NQYQGNGYQQQQGQSGGAFSSLA 271

Qy 55 -----YQGQYQOQYNPDAQYQQYQNPQG---GYQYQNPQG-----GYQYQNPQG----- 96
Db 272 SMAQSYLGGQQTN----QGYQYQGQNNQYQOGONQYQHQQGQQQD----QCHSS 323

S12921

suppressor 2 protein - yeast (*Pichia Pinus*)

N:Alternative names: Sup2 protein

C:Species: *Pichia Pinus*

C:Accession: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 02-Feb-2001

R;Kushnirov, V.V.; Ter-Avanesyan, M.D.; Didichenko, S.A.; Smirnov, V.N.; Chernoff, Y.

Yeast 6, 461-472, 1990

A:title: Divergence and conservation of sup2(sup35) gene of yeasts *Pichia pinus* and *S.*

A:Reference number: S12921; MUID:91161341

A:Accession: S12921

A:Molecule type: DNA

A:Residues: 1-741 <KUS>

A:Cross-references: EMBL:X56910; NID:93235; PIDN:CAA40231.1; PID:93236

A:Gene: SUP2

C:Superfamily: GTP binding; nucleotide binding; P-loop; phosphoprotein

C:Keywords: GTP binding; nucleotide-binding factor Tu homology <ERU>

F;319-467/Region: nucleotide-binding motif A (P-loop)

F;225-312/Region: GTP-binding NKXD motif

F;331/Binding site: GTP (Lys) #status predicted

RESULT 4

RNA/ssDNA-binding protein TAFI168 - human

N:Alternative names: TATA-binding protein-associated factor TAFI168

C:Species: Homo sapiens (man)

C:Accession: S17954 #sequence_revision 17-Apr-1998 #text_change 20-Jun-2000

R:Bertolotti, A.; Lutz, Y.; Heard, D.J.; Chamblon, P.; Tora, L.

EMBO J. 15, 5022-5031, 1996

A:Title: HTAFI168, a novel RNA/ssDNA-binding protein with homology to the pro-oncoprotein TAFII168.

A:Reference number: S17954; MUID:97045110

A:Accession: S17954

A:Molecule type: mRNA

A:Residues: 1-589 <BER>

A:Cross-references: EMBL:X98993; NID:91628402; PIDN:CAA67398.1; PID:91628403

A:Experimental source: HeLa cells

C:Genetics:

A:Gene: taf68

C:Complex: up to 13 TATA-binding protein-associated factors (TAFIIs) form together w/1

C:Function:

A:Description: required for activated transcription; binds RNA and single stranded DNA

C:Superfamily: RNA-binding protein EWS; ribonucleoprotein repeat homology

C:Keywords: RNA binding

F;232-307/Domain: ribonucleoprotein repeat homology <RRM>

Query Match 23.3%; Score 163.5; DB 2; Length 589;

Best Local Similarity 30.8%; Pred. No. 3.8e-07; Mismatches 49; Conservative 13; Indels 55; Gaps 8;

A:Reference number: S12921; MUID:97045110

C:Species: *Pichia Pinus*

C:Accession: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 02-Feb-2001

R;Kushnirov, V.V.; Ter-Avanesyan, M.D.; Didichenko, S.A.; Smirnov, V.N.; Chernoff, Y.

Yeast 6, 461-472, 1990

A:title: Divergence and conservation of sup2(sup35) gene of yeasts *Pichia pinus* and *S.*

A:Reference number: S12921; MUID:91161341

A:Accession: S12921

A:Molecule type: DNA

A:Residues: 1-741 <KUS>

A:Cross-references: EMBL:X56910; NID:93235; PIDN:CAA40231.1; PID:93236

A:Gene: SUP2

C:Superfamily: GTP binding; nucleotide binding; P-loop; phosphoprotein

C:Keywords: GTP binding; nucleotide-binding factor Tu homology <ERU>

F;225-312/Region: nucleotide-binding NKXD motif

F;331/Binding site: GTP (Lys) #status predicted

Query Match 23.7%; Score 166; DB 2; Length 405;

Best Local Similarity 34.4%; Pred. No. 1.6e-07; Mismatches 55; Conservative 7; Indels 54; Gaps 9;

A:Reference number: A93093; MUID:8728684

A:Accession: C27061

A: Molecule type: DNA

A:Residues: 1-180, 'T', 182-322, 'VHSOLNLPWQVPTWAITPIQIRVMGA NR LMSMDRN', 'RMVNSNPMS TEDRNTAE'

A:Cross-references: EMBL:M16717; NID:9171534; PIDN:AA34615.1; PID:9171536

C:Genetics:
A:Map position: 3L

RESULT 3

RESLT 3

Db	57 GOSYSOSYGGYENQKQSSYSSQQPNNNQGQQONMSESSSGSGQGRAPSYDQPDYQODDSYDQ 116	Qy	57 QGYQQINPOGGYQOCFPNPOGGRGYKKNFNYNNNITQGYQA 115	Db	111 OSYGQO-----AFAGQORPQOGFQNGFSQGMSNSRGQGQYQONQGFGQGSFS 165
Qy	77 GGYQQINPOGGYQOCFPNPOGGRGYKKNFNYNNNITQGYQA 115	Db	117 SGYDQH--QGSYEQDSN----YDQHQDSYISQNQOSYHS 148	Qy	103 -----NEYNNNLYQYAGFQDQSQ 123
Db	117 SGYDQH--QGSYEQDSN----YDQHQDSYISQNQOSYHS 148	Db	166 GQSGWGSNSLSSANSNGNNN-QGSSSGYQ-NNQG 197	Qy	57 QGYQQINPOGGYQOCFPNPOGGRGYKKNFNYNNNITQGYQA 115
RESULT	5	RESULT	7	Qy	57 QGYQQINPOGGYQOCFPNPOGGRGYKKNFNYNNNITQGYQA 115
T04866	DAG proteo... homolog F18F4.120 - <i>Arabidopsis thaliana</i>	C;Species: <i>Arabidopsis thaliana</i> (mouse-ear cress)	T21998	Db	111 OSYGQO-----AFAGQORPQOGFQNGFSQGMSNSRGQGQYQONQGFGQGSFS 165
C;Date: 23-Apr-1999	#sequence_revision 23-Apr-1999	C;Accession: T04866	C;Species: Caenorhabditis elegans	Qy	103 -----NEYNNNLYQYAGFQDQSQ 123
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.	submitted to the Protein Sequence Database, February 1998	R;McMurtry, A.	C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999	Db	166 GQSGWGSNSLSSANSNGNNN-QGSSSGYQ-NNQG 197
A;Reference number: 215388	A;Accession: T04866	R;Ainscough, R.	C;Accession: T21998	Qy	57 QGYQQINPOGGYQOCFPNPOGGRGYKKNFNYNNNITQGYQA 115
A;Molecule type: DNA	A;Residues: 1-419 <BEV>	A;Reference number: Z19499	A;Accession: T21998	Db	111 OSYGQO-----AFAGQORPQOGFQNGFSQGMSNSRGQGQYQONQGFGQGSFS 165
A;Cross-references: EMBL:AL021637	A;Experimental source: cultivar Columbia; BAC clone F18F4	A;Status: Preliminary; translated from GB/EMBL/DDJB	A;Residues: 1-945 <WIL>	Qy	103 -----NEYNNNLYQYAGFQDQSQ 123
A;Genetics:	A;Map position: 4	A;Cross-references: EMBL:Z69791; PIDN:CAA93660.1; GSPDB:GN00028; CESP:F39D8.1a	A;Experimental source: clone F39D8	Db	166 GQSGWGSNSLSSANSNGNNN-QGSSSGYQ-NNQG 197
A;Introns: 123/2; 156/1; 178/1	A;Note: F18F4.120	C;Genetics:	C;Gene: CESP:F39D8.1a	Qy	57 QGYQQINPOGGYQOCFPNPOGGRGYKKNFNYNNNITQGYQA 115
Query Match	21.1%; Score 148; DB 2; Length 419;	Best Local Similarity	34.8%; Pred. No. 6.1e-06;	Qy	57 QGYQQINPOGGYQOCFPNPOGGRGYKKNFNYNNNITQGYQA 115
Matches	57; Conservative	Matches	9; Mismatches	Db	111 OSYGQO-----AFAGQORPQOGFQNGFSQGMSNSRGQGQYQONQGFGQGSFS 165
Qy	11 ONYQOQYSQNGNQ---QGQNNRY---QGY---QANQNAQOP---AGGYQON-----YQG 51	Db	215 QNFQNRNTQYQGQOPPMQGGGSYGPQOGYATPGQGQGTQAPPFFQOGYNQOPRSPPPPYQ- 273	Qy	1 MSDSNQGNQNNQNYQYSQNGNQGQGNRNYQGYQYQANA-QAQ---PAGGYQONY-OGYSGV 55
Db	274 YSGCQG-----GQYQYN-----PDAGYQCYQINPOGGYQOYN-POGGIQQ--- 90	Db	227 VGDQNQNOFQDPNQSGYQYQONQNIINTGGYQIDOSQONQONQSGYQDPDNGQYQONQONQNI 286	Qy	56 QGGTQOYNPDRAGYQOQYNPQG---GQYQYNP-----OGGYQOQEN---POGRGNY 101
Qy	91 QFNPOGGGRGYNKFNFWNNNL-----QGYQAGFOPOSQ 122	Db	287 NAGGYVQDNQFQGQDQINAGGYVQYQONQNGQFQDQDNQNYQGYQGSTLAPSAGEVDY 346	Db	332 NYNDQGA-GNFGPASAGAGNLGPAPQAGNPNQIGQGYSGPGEQNQ 374
Db	347 QPFPSGNTDINNNAOGFDQDQGQGQYQ 375	Query Match	21.0%; Score 147.5; DB 2; Length 945;	Qy	102 K-----NFNNNNNLOGY---QAGFOPOSQ 122
RESULT	6	Best Local Similarity	34.9%; Pred. No. 1.6e-05;	Db	347 QPFPSGNTDINNNAOGFDQDQGQGQYQ 375
T20847	hypothetical protein F13E9.4 - <i>Caenorhabditis elegans</i>	Matches	52; Conservative	Qy	102 K-----NFNNNNNLOGY---QAGFOPOSQ 122
C;Species: <i>Caenorhabditis elegans</i>	C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999	Matches	7; Mismatches	Db	347 QPFPSGNTDINNNAOGFDQDQGQGQYQ 375
C;Accession: T20847	R;McMurtry, A.	Query Match	21.0%; Score 147.5; DB 2; Length 945;	Qy	102 K-----NFNNNNNLOGY---QAGFOPOSQ 122
submitted to the EMBL Data Library, February 1996	A;Reference number: Z19332	Best Local Similarity	34.9%; Pred. No. 1.6e-05;	Db	347 QPFPSGNTDINNNAOGFDQDQGQGQYQ 375
A;Accession: T20847	A;Status: preliminary; translated from GB/EMBL/DDJB	Matches	52; Conservative	Qy	102 K-----NFNNNNNLOGY---QAGFOPOSQ 122
A;Structure: preliminary; translated from GB/EMBL/DDJB	A;Molecule type: DNA	Matches	7; Mismatches	Db	347 QPFPSGNTDINNNAOGFDQDQGQGQYQ 375
A;Residues: 1-409 <WIL>	A;Cross-references: EMBL:Z69383; PIDN:CAA93412.1; GSPDB:GN00022; CESP:F13E9.4	Query Match	21.0%; Score 147.5; DB 2; Length 409;	Qy	102 K-----NFNNNNNLOGY---QAGFOPOSQ 122
A;Experimental source: clone F13E9	A;Genetics:	Best Local Similarity	31.2%; Pred. No. 6.e-06;	Db	347 QPFPSGNTDINNNAOGFDQDQGQGQYQ 375
A;Introns: 32/1; 275/3; 337/3	A;Note: F13E9.4	Matches	15; Mismatches	Qy	102 K-----NFNNNNNLOGY---QAGFOPOSQ 122
Query Match	21.0%; Score 147.5; DB 2; Length 409;	Query Match	21.0%; Score 147.5; DB 2; Length 409;	Db	347 QPFPSGNTDINNNAOGFDQDQGQGQYQ 375
Best Local Similarity	31.2%; Pred. No. 6.e-06;	Best Local Similarity	31.2%; Pred. No. 6.e-06;	Qy	102 K-----NFNNNNNLOGY---QAGFOPOSQ 122
Matches	48; Conservative	Matches	15; Mismatches	Db	347 QPFPSGNTDINNNAOGFDQDQGQGQYQ 375
R;Doering, V.; Schleicher, M.; Noedel, A.A.	J. Biol. Chem. 266, 17509-17515, 1991	R;Doering, V.; Schleicher, M.; Noedel, A.A.	J. Biol. Chem. 266, 17509-17515, 1991	Qy	102 K-----NFNNNNNLOGY---QAGFOPOSQ 122
Qy	4 SNGQGNQNYQO-YSONGQNOQQNRYQGYQAYNAQAPAGGYQON-----YQGYQYQ 56	A;Title: Dictyostelium annexin VII (synexin). cDNA sequence and isolation of a gene d	A;Title: Dictyostelium annexin VII (synexin). cDNA sequence and isolation of a gene d	Db	347 QPFPSGNTDINNNAOGFDQDQGQGQYQ 375
Db	55 TQGQMLESSYQOGQGQNGSMQGSS---QGYGCGNSQDQGYGSQSGSOMGVOQYGGSS 110	A;Accession: A40977	A;Accession: A40977	Qy	102 K-----NFNNNNNLOGY---QAGFOPOSQ 122
A;Molecule type: mRNA	A;Accession: S22756	A;Molecule type: mRNA	A;Accession: S22756	Db	347 QPFPSGNTDINNNAOGFDQDQGQGQYQ 375
A;Residues: 17-462 <GRE2>	A;Residues: 1-39 <GRE1>	A;Residues: 17-462 <GRE2>	A;Residues: 17-462 <GRE2>	Qy	102 K-----NFNNNNNLOGY---QAGFOPOSQ 122
A;Cross-references: EMBL:X60270; PIDN:97209; PIDN:CAA42816.1; PID:9671859	A;Cross-references: EMBL:X60270; PIDN:97209; PIDN:CAA42816.1; PID:9671859	A;Cross-references: EMBL:X60270; PIDN:97209; PIDN:CAA42816.1; PID:9671859	A;Cross-references: EMBL:X60270; PIDN:97209; PIDN:CAA42816.1; PID:9671859	Db	347 QPFPSGNTDINNNAOGFDQDQGQGQYQ 375
A;Reference number: A40977; MUID:9137375	A;Accession: A40977	A;Reference number: A40977; MUID:9137375	A;Reference number: A40977; MUID:9137375	Qy	102 K-----NFNNNNNLOGY---QAGFOPOSQ 122
A;Accession: A40977	A;Accession: S22756	A;Accession: S22756	A;Accession: S22756	Db	347 QPFPSGNTDINNNAOGFDQDQGQGQYQ 375
A;Molecule type: mRNA	A;Accession: S22756	A;Accession: S22756	A;Accession: S22756	Qy	102 K-----NFNNNNNLOGY---QAGFOPOSQ 122

Query Match 20.3%; Score 142; DB 2; Length 676;
 Best Local Similarity 36.6%; Pred. No. 3.4e-05; Mismatches 48; Conservative 2; Indels 32; Gaps 7; Matches 48; Mismatches 49; Indels 32; Gaps 7;

Qy 8 NNOQNQYQSQNQNO-QGNRNYQQYQAYNAQAQPAGGYQNY----QGYSQY-QGG 59
 Db 230 NQQPSYGGYQOPPQPGYQGNSYDPPYQGPOQQPYPGGGQOPPYQGNSNQGGGGYQPGQG 289

Qy 60 YQQYNPDAGYQQ--QYNN-----POGGY---QOYNPQGGYQQQFNPOGGGRNYK 102
 Db 290 APYPCGSGGYQPAQYQFSGQGSAPQPNQGGYQPPQQQYQGGQGQGGQGQGGFFP 349

Qy 103 NFNYYNNLLOGY 113
 Db 350 N-----QGY 353

RESULT 12

S46269
 synonymous sarcoma translocation protein - human
 N;Alternate names: SYT protein
 C;Species: Homo sapiens (man)
 C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 08-Oct-1999
 C;Accession: S46269; S51172
 Nature Genet. 7, 502-508, 1994
 A;Title: Identification of novel genes, SYT and SSX, involved in the t(X;18) (p11.2;q11.1)
 A;Reference number: S46269; MUID:95038836
 A;Molecule type: mRNA
 A;Residues: 1404 <CL>
 A;Cross-references: EMBL:X79201
 A;Note: it is uncertain whether translation is initiated upstream to the sequenced region
 R;Cooper, C.S.
 R;Submitted to the EMBL Data Library, May 1994
 A;Reference number: S51172
 A;Accession: S51172
 A;Molecule type: mRNA
 A;Residues: 1404 <COO>
 A;Cross-references: EMBL:X79201; NID:9531105; PIDN:CAA55792.1; PID:9531106
 C;Genetics:
 A;Gene: GDB:SSY; SYT
 A;Cross-references: GDB:434322; OMIM:600192
 A;Map position: 18q11.2-18q11.2

Query Match 20.1%; Score 141; DB 2; Length 404;
 Best Local Similarity 35.7%; Pred. No. 2.4e-05; Mismatches 50; Conservative 9; Indels 55; Gaps 10; Matches 50; Mismatches 55; Indels 26; Gaps 10;

Qy 1 MSDSOGNNNQNYQY---SONGNOQ-GNNRQGY-AYNAAQAPAG--GYQNY 49
 Db 252 MGQVNQNHMGQRPYRPPQPPQYQSGQEDYQHGSGQGPBGMNQYYPDG 311

Qy 50 QGYQSQQGQYQYQYDAGY---QOQYNPQGGYQQYNPQGGYQQQFNP-QGGRG-NYKNF 104
 Db 312 NSQVTCQGQDQAYQPPQPPQGQQQYQPGQGQY---GQOOGYGPSCQGPQYQPNY 365

Qy 105 NYNNILQGYO-AGRPQSQG 123
 Db 366 POG---QGQYGGYRPTQG 382

RESULT 13

151653 dsRNA-binding protein 4F 2 - African clawed frog (fragment)
 C;Species: Xenopus laevis (African clawed frog)
 C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C;Accession: 151653
 R;Babs, B.L.; Hurst, S.R.; Singer, J.D.
 A;Title: Binding properties of newly identified Xenopus proteins containing dsRNA-bindin
 A;Reference number: 151652; MUID:95006648

Query Match 20.3%; Score 142; DB 2; Length 676;
 Best Local Similarity 36.6%; Pred. No. 3.4e-05; Mismatches 48; Conservative 2; Indels 32; Gaps 7; Matches 48; Mismatches 49; Indels 32; Gaps 7;

Qy 7 GNNQO----NYQQYSQNGNQ----QCNRNYQY---QAYNA-----
 Db 633 QQQQQRPPQHASNPPKPSYQSYQGQH-QGGQQQQPQQQQQTYN-QNQYSNYGPQKQ 690

Qy 86 GGYQQFNPNQGGRNYKFNYYNNLLOGY 112
 Db 691 GGYINOQTOGAASAGSY---NISNSYNG 74

RESULT 14

T18240 hypothetical transmembrane protein - yeast (Candida albicans)
 C;Species: Candida albicans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999
 C;Accession: T18240
 R;Barrell, B.G.; Ralbandream, M.A.
 Submitted to the EMBL Data Library, November 1998
 A;Reference number: 218831
 A;Accession: T18240
 A;Status: Preliminary; translated from GB/EMBL/DDJB
 A;Molecule type: DNA
 A;Residues: 1-236 <BAR>
 A;Cross-references: EMBL:AL033503; NID:e1341066; PID:e1341067; PIDN:CAA22011.1
 C;Genetics:
 A;Note: Caa9C4.01
 C;Superfamily: WW repeat homology <WWR>
 F;10-48/Domain: WW repeat homology <WWR>

Query Match 20.0%; Score 140.5; DB 2; Length 236;
 Best Local Similarity 31.9%; Pred. No. 1.5e-05; Mismatches 45; Conservative 12; Indels 39; Gaps 6; Matches 45; Mismatches 45; Indels 39; Gaps 6;

Qy 8 NNOQNQO----QSQNQ---NOQQGNRNYQYQAYNAQAQPAGGYQNYQGSGYQ 56
 Db 86 NWQRDXQGPQGOVYQGPQGQVQGPQPSQHQY---QGPQGQYQ---QGPP 136

Qy 57 QGYQQYNPDAG-YQQYNPQGGYQQYNPQGGYQQQFNP-QGGRG-NYKNF 100
 Db 137 PGYYQGPSPQSYQYQGPQPSQYQYQGPPQPSQYQYQGPPQPSQYQYQGPPQPSQYQYQ 196

Qy 101 --YKFNYYNNLLOGYQAGFO 118
 Db 197 GNAINSWEDHERMEGYQDGQ 217

RESULT 15

T02627 hypothetical protein At2g25960 [imported] - Arabidopsis thaliana
 N;Alternate names: hypothetical protein T19L18.23
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Feb-2001
 C;Accession: T02627; HB4654
 R;Rounseley, S.D.; Kaul, S.; Lin, X.; Kettchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke
 Submitted to the EMBL Data Library, August 1998
 A;Description: Arabidopsis thaliana chromosome II BAC T19L18 genomic sequence.

A; Reference number: 214681
 A; Accession: T02627
 A; Status: translated from GB/EMBL/DBJ
 A; Molecule type: DNA
 A; Residues: 1-632 <ROUS>
 A; Cross-references: EMBL:AC004747; NID:93413696; PID:93413716
 A; Experimental source: cultivar Columbia
 R, Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, J.;
 Goss, D.; Neerman, W.C.; White, O.; Elsen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.;
Nature 402, 661-668, 1999
 A; Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A; Reference number: AB4420; MUID:20083487
 A; Accession: H84654
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-632 <STO>
 A; Cross-references: GB:AE002093; NID:93413716; PIDN: AAC31239.1; GSPDB:GN00139
 C; Genetics:
 A; Gene: T19L18.23; At2g25960
 A; Map position: 2
 A; Introns: 112/1; 206/3; 240/3; 264/3; 304/3
 C; Superfamily: glutinin

Query Match 19.8%; Score 139; DB 2; Length 632;
 Best local Similarity 31.7%; Pred. No. 5.8e-05; Gaps 7;
 Matches 46; Conservative 6; Mismatches 53; Indels 40;

Qy	13	YQQYSQNGNQQCNRYYQGYQAVANAOQAPAGG-.....YONYQGY----SGYQ 56
		: : : : : : : : : : : :
Db	374	YDQSSVPPSQSAQEQDYGGQQSQPSGGSSAPPDTIGNYQHASGYQQAGQYQ 433
Qy	57	QGYQQNPAGQQYQQYNPQGYQQYQWPGQYQQQFNP-----QGGGRNY 101
Db	434	QGYGAYN-ASCGSGQGAAGDQ--QGGGSTTNPSEEDASQAPPSSAQSGGQY 488
Qy	102	KNNNNNNLQGY--QAGE--QQSQ 122
Db	489	GTTGQQPPAQGSQGQAGYGGAPPSQ 513

Search completed: February 1, 2002, 15:02:47
 Job time: 217 sec

Sequence 4, Appl
Sequence 1, Appl
Sequence 4, Appl
Sequence 18, Appl
Sequence 91, Appl
Sequence 2, Appl
Sequence 61, Appl
Sequence 10, Appl
Sequence 4, Appl
Sequence 6, Appl
Sequence 5, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 1, Appl
Sequence 1, Appl

NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESS: NIKaido, Marmelstein, Murray & Oram
 STREET: 655 Fifteenth Street N.W. Suite 330
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005-5701
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/982.112
 FILING DATE: 1992/11/25
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US/07/617,102
 FILING DATE: 21-NOV-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Chin, Monica F.
 REGISTRATION NUMBER: 36-105
 REFERENCE/DOCKET NUMBER: 564-1906
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 638-5000
 TELEFAX: (202) 638-4810
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 362 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-982-112-2

RESULT 5
 US-08-437-027-21
 ; Sequence 21, Application US/08437027
 ; Patent No. 5670317
 ; GENERAL INFORMATION:
 ; APPLICANT: Landayi, Marc
 ; APPLICANT: Gerald, William
 ; TITLE OF INVENTION: A DIAGNOSTIC TEST FOR TEST FOR THE DESMOPLASTIC
 ; TITLE OF INVENTION: SMALL ROUND CELL TUMOR
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Cooper & Dunham LLP
 ; STREET: 1185 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10035
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/437, 027
 ; FILING DATE: 536
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: White, John P
 ; REGISTRATION NUMBER: 28,678
 ; REFERENCE/DOCKET NUMBER: 46416/JPW/CCA
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-278-0400
 ; TELEFAX: 212-391-0525
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 365 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGI: unknown
 ; MOLECULE TYPE: protein
 ; US-08-437-027-20

Query Match 14.8%; Score 103.5; DB 1; Length 365;
 Best Local Similarity 31.8%; Pred. No. 0.011; Indels 23; Gaps 4;
 Matches 35; Conservative 4; Mismatches 48; Indels 23; Gaps 4;

RESULT 7
 US 08-343-443B-2
 Sequence 2, Application US/08343443B
 ;
 GENERAL INFORMATION:
 ; Patent No. 5968734
 ;
 APPLICANT: Aurias, Alain
 APPLICANT: Delatre, Olivier
 APPLICANT: Desmae, Chantal
 APPLICANT: Melot, Thomas
 APPLICANT: Peter, Martine
 APPLICANT: Ploouastel, Beatrice
 APPLICANT: Thomas, Gilles
 APPLICANT: Zucman, Jessica
 TITLE OF INVENTION: NUCLEIC ACID CORRESPONDING TO A GENE OF CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL TRANSLATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID TRANSLOCATIONS
 NUMBER OF SEQUENCES: 129
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Weiser & Associates
 STREET: 210 South Fifteenth Street
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19102
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: AEDIT 1.0 DOS text editor
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/343,443B
 FILING DATE: 10-NOV-1994
 CLASSIFICATION: 514
 PRIORITY APPLICATION NUMBER: PCT/FR93/00494
 FILING DATE: 19-MAY-1993
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: FR 92/06123
 FILING DATE: 20-MAY-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Weiser, Gerard J.
 REFERENCE/DOCKET NUMBER: 19,763
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-875-8383
 TELEFAX: 215-875-8394
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 656 amino acids
 TYPE: amino acid
 MOLECULE TYPE: protein
 US-08-343-443B-2

Query Match 14.8%; Score 103.5; DB 4; Length 656;
 Best Local Similarity 31.8%; Pred. No. 0.011; Indels 23; Gaps 4;
 Matches 35; Conservative 4; Mismatches 48; Indels 23; Gaps 4;

RESULT 8
 US-09-214-564A-4
 Sequence 4, Application US/09214564A
 ;
 Patent No. 6150515
 ;
 GENERAL INFORMATION:
 ;
 APPLICANT: Sharp, Phillip A.
 APPLICANT: Zhou, Qiang
 TITLE OF INVENTION: TAT-SF: Cofactor For Stimulation Of Transcriptional Elongation By HIV-1 TAT
 FILE REFERENCE: M0656/7042
 CURRENT APPLICATION NUMBER: US/09/214,564A
 CURRENT FILING DATE: 1998-08-18
 PRIORITY APPLICATION NUMBER: US 60/021,218
 PRIORITY FILING DATE: 1996-07-03
 PRIORITY APPLICATION NUMBER: US 60/033,152
 PRIORITY FILING DATE: 1996-12-13
 PRIORITY APPLICATION NUMBER: PCT/US97/11713
 PRIORITY FILING DATE: 1997-07-03
 NUMBER OF SEQ ID NOS: 5
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 4
 LENGTH: 656
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-214-564A-4

Query Match 14.8%; Score 103.5; DB 4; Length 656;
 Best Local Similarity 31.8%; Pred. No. 0.011; Indels 23; Gaps 4;
 Matches 35; Conservative 4; Mismatches 48; Indels 23; Gaps 4;

RESULT 9
 US-08-536-711A-8
 Sequence 8, Application US/08538711A
 ;
 Patent No. 5994052
 ;
 GENERAL INFORMATION:
 ;
 APPLICANT: Mulshine, James, L.
 TITLE OF INVENTION: AN EPITHELIAL PROTEIN AND DNA THEREOF FOR USE IN EARLY CANCER DETECTION
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
 STREET: 345 PARK AVENUE
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10154
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/538,711A
 FILING DATE: 02-OCT-1995
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: KATHRYN M. BROWN
 REGISTRATION NUMBER: 34,556
 REFERENCE/DOCKET NUMBER: 2026-4201
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 758-4800
 TELEFAX: (212) 751-6849
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 341
 TYPE: Amino Acid
 STRANDEDNESS: Unknown
 TOPOLOGY: Linear
 MOLECULE TYPE: Peptide
 US-08-725-027-8

RESULT 10

Query Match 14.7%; Score 103; DB 2; Length 341;

Best Local Similarity 33.3%; Pred. No. 0.0054; Matches 33; Conservative 6; Mismatches 44; Indels 16; Gaps 6;

Db 216 RGGSGCGGSGRGFGGGYNGG--GGPGGGNFGGGSPGYY-GRRGGGGGGPGY---NO 267

Qy 76 QGGYQQYNPOGGYQQQFNPOGGGRGYKNF-NYNNNLQGY 113

Db 268 GGGY----GGYDNYGGGNYGSGNYNDFGNYNQOPSNY 301

RESULT 11

Sequence 7, Application US/08538711A

PATENT NO. 5984062

GENERAL INFORMATION:

APPLICANT: MULSHINE, JAMES, L.

TITLE OF INVENTION: AN EPITHELIAL PROTEIN AND

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L. L. P.

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/538,711A

FILING DATE: 02-OCT-1995

CLASSIFICATION: 435

PRIORITY APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: KATHRYN M. BROWN

REGISTRATION NUMBER: 34,556

REFERENCE/DOCKET NUMBER: 2026-4201

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 341

TYPE: Amino Acid

STRANDEDNESS: Unknown

TOPOLOGY: Linear

MOLECULE TYPE: Peptide

US-08-538-711A-7

Query Match 14.7%; Score 103; DB 2; Length 353;

Best Local Similarity 33.3%; Pred. No. 0.0056; Matches 33; Conservative 6; Mismatches 44; Indels 16; Gaps 6;

Qy 18 QNGNQQQGNR--YGGYQAVNAQAQPAGGYQNYQGYSQGQGGYQOQNPDAGYQOQNP 75

Db 216 RGGSGCGGSGRGFGGGYNGG--GGPGGGNFGGGSPGYY-GRRGGGGGGPGY---NO 267

Qy 76 QGGYQQYNPOGGYQQQFNPOGGGRGYKNF-NYNNNLQGY 113

Db 268 GGGY----GGYDNYGGGNYGSGNYNDFGNYNQOPSNY 301

Query Match 14.7%; Score 103; DB 2; Length 353;

Best Local Similarity 33.3%; Pred. No. 0.0056; Matches 33; Conservative 6; Mismatches 44; Indels 16; Gaps 6;

Qy 18 QNGNQQQGNR--YGGYQAVNAQAQPAGGYQNYQGYSQGQGGYQOQNPDAGYQOQNP 75

Db 216 RGGSGCGGSGRGFGGGYNGG--GGPGGGNFGGGSPGYY-GRRGGGGGGPGY---NO 267

Qy 76 QGGYQQYNPOGGYQQQFNPOGGGRGYKNF-NYNNNLQGY 113

Db 268 GGGY----GGYDNYGGGNYGSGNYNDFGNYNQOPSNY 301

APPLICANT: Brown, Steven
 APPLICANT: Kellett, Patti
 APPLICANT: Essigmann, John M.
 APPLICANT: Lippard, Stephen J.
 TITLE OF INVENTION: DNA Structure Specific Recognition
 TITLE OF INVENTION: Protein and Uses Therefor
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 STREET: 2 Militia Drive
 CITY: Lexington
 STATE: MA
 COUNTRY: USA
 ZIP: 02173
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/258,442
 FILING DATE:
 CLASSIFICATION: 530
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 07/539,906
 FILING DATE: 16-JUN-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Granahan, Patricia
 REGISTRATION NUMBER: 32,227
 REFERENCE/DOCKET NUMBER: MTR-4787AAA
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-861-6240
 TELEFAX: 617-861-9540
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 542 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: *Saccharomyces cerevisiae*
 IMMEDIATE SOURCE:
 CLONE: fractional yeast SSRP (fySSRP) (predicted)
 ; US-08-258-442-13

Query Match 13.3%; Score 93; DB 1; Length 542;
 Best Local Similarity 27.9%; Pred. No. 0.092; Gaps 5;
 Matches 38; Conservative 5; Mismatches 59; Indels 34; Gaps 5;

```

QY 6 QGNNNQQNYQQYSQNQNGNQQGNNRNYQQYQAVNAQAGPAGGYV-----ONYQ 50
Db 5 QAOQPQQQQQQQQQQQQQQQA--PYQGHFQQSPQQQQNVFPPLPQLQTSQSQQQQ 62
QY 51 GYSGYQQGGYQQYN----PDAGYQQOYNPOGGYQQYNPOGG----YQOOFNP 94
Db 63 QYANSNSNSNNNNNNNALPQDFGYMQOTGSQNYPTINQQOFSEFYNSFLSHLTQTNP 122
QY 95 Q-GGGGNYKHFYNNN 109
Db 123 SVTGTGASSNNNNNN 138

```

RESULT 15
 US-08-328-809-8
 ; Sequence 8, Application US/08328809
 ; Patent No. 5705334
 GENERAL INFORMATION:
 APPLICANT: Lippard, Stephen J.
 APPLICANT: Essigmann, John M.
 APPLICANT: Donahue, Brian A.
 APPLICANT: Toney, Jeffrey H.
 APPLICANT: Bruhn, Suzanne L.

APPLICANT: Pil, Piefer M.
 APPLICANT: Brown, Steven
 APPLICANT: Kellett, Patti
 TITLE OF INVENTION: Uses For DNA Structure-Specific
 TITLE OF INVENTION: Recognition Proteins
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Patent Administrator, Testa, Hurwitz & Thibeault
 STREET: 53 State Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/328,809
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Fenton, Gillian M.
 REGISTRATION NUMBER: 36,508
 REFERENCE/DOCKET NUMBER: MTR-023 (5473/24)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-248-7000
 TELEFAX: 617-248-7100
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 542 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 IMMEDIATE SOURCE:
 CLONE: fractional yeast SSRP (fySSRP) (predicted)
 ; US-08-328-809-8

Query Match 13.3%; Score 93; DB 1; Length 542;
 Best Local Similarity 27.9%; Pred. No. 0.092; Gaps 5;
 Matches 38; Conservative 5; Mismatches 59; Indels 34; Gaps 5;

```

QY 6 QGNNNQQNYQQYSQNQNGNQQGNNRNYQQYQAVNAQAGPAGGYV-----ONYQ 50
Db 5 QAOQPQQQQQQQQQQQQQA--PYQGHFQQSPQQQQNVFPPLPQLQTSQSQQQQ 62
QY 51 GYSGYQQGGYQQYN----PDAGYQQOYNPOGGYQQYNPOGG----YQOOFNP 94
Db 63 QYANSNSNSNNNNNNNALPQDFGYMQOTGSQNYPTINQQOFSEFYNSFLSHLTQTNP 122
QY 95 Q-GGGGNYKHFYNNN 109
Db 123 SVTGTGASSNNNNNN 138

```

Search completed: February 1, 2002, 15:02:16
 Job time: 236 sec

GanCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - Protein search, using sw model

Run on:

February 1, 2002, 14:57:35 ; Search time 38.77 Seconds

(without alignments)
235.002 Million cell updates/sec

Title: US-09-591-632-2_COPY_1_123

Perfect score: 701

Sequence: 1 MSDSNQGNQQNQQIQSQNG FNYNNNLQGYQAGFQPOSQG 123

Scoring table: BLASTM62

Gapext 10.0 , Gapext 0.5

Searched: 522463 seqs, 7407290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq-1101:*

1: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:980.DAT: *

2: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:981.DAT: *

3: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:982.DAT: *

4: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:983.DAT: *

5: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:984.DAT: *

6: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:985.DAT: *

7: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:986.DAT: *

8: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:987.DAT: *

9: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:988.DAT: *

10: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:989.DAT: *

11: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:990.DAT: *

12: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:991.DAT: *

13: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:992.DAT: *

14: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:993.DAT: *

15: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:994.DAT: *

16: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:995.DAT: *

17: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:996.DAT: *

18: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:997.DAT: *

19: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:998.DAT: *

20: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA1:999.DAT: *

21: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA2:000.DAT: *

22: /SIDS8/gcdata/geneseq/geneseq/geneseq/AA2:001.DAT: *

ALIGNMENTS	
RESULT 1	ID AAB30792 standard; Protein: 685 AA.
ID XX	ID AAB30792 standard; Protein: 685 AA.
AC XX	AC AAB30792;
DT XX	DT 02-APR-2001 (first entry)
DE XX	DE Amino acid sequence of a yeast Sup35 protein.
XX	XX Sup35; phenotype: SChG; self-coalesce; higher-order aggregate; amyloidogenic domain; aggregation; fibril; phenotypic alteration; gene therapy; disease resistance; plant pigmentation; prion disease.
KW XX	KW amino acid sequence; gene therapy; disease resistance; plant pigmentation; prion disease.
OS XX	OS Saccharomyces cerevisiae.
XX	XX
PN XX	PN WO200075324-A2.
PD XX	PD 14-DEC-2000.
PR XX	PR 09-JUN-2000; 2000WO-US15876.
PA XX	PA (ARCH-) ARCH DEV CORP.
PI XX	PI Lindquist S, Li L, Ma J, Liu J, Sondheimer N, Scheibel T;
XX	XX DR WPI; 2001-051723-077.
DR XX	DR N-PSDB; AAC86679.
PT XX	PT New nucleic acid encoding chimeric proteins with self-assembly properties, useful e.g. for diagnosis and treatment of prion diseases, also related aggregates, fibrils and polymers.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	701	100.0	685	22	AAB30792
2	682	97.3	271	22	AAB30800
3	450	64.2	215	22	AAB30799
4	228	32.5	715	22	AAB30820
5	166	23.7	405	22	AAB30810
6	166	23.7	405	22	AAB30821
7	166	23.7	741	22	AAB30819
8	160	22.8	407	10	AAP94879
9	148	21.1	417	21	AGC30013
10	148	21.1	419	21	AGC30012
11	148	21.1	439	21	AAG30011

xx
PS
Claim 11: Page 125-127: 100pp: English

Claim 11: Page 125-127: 188pp: English

PT New nucleic acid encoding chimeric proteins with self-assembly properties, useful e.g. for diagnosis and treatment of prion diseases.

XX

also related aggregates, fibrils and p

polymers -

The present sequence represents a yeast Sup35 protein. The protein possesses the prion-like capacity to undergo a self-perpetuating conformational alteration that changes the functional state of Sup35 in a manner that creates a heritable change in phenotype. It is used to construct chimeric polypeptides of the invention, which comprise at least one SCHAG (self-coalesces into higher-order aggregates) amino acid sequence fused in frame with a polypeptide of interest (which is other than a marker protein, a glutathione-S-transferase or a staphylococcal nuclear protein). The specification also describes chimeric polypeptides that comprises an amyloidogenic domain that causes aggregation into fibrils. The chimeric polypeptides are used to prepare polymers with multiple reactivities, e.g., derivatised with enzymes, or specific binding partners, and useful e.g. for performing multi-step chemical reactions. They can be used to create an inducible, or stable phenotypic alteration in a cell, e.g. for gene therapy, protein production, imparting disease resistance to plants, altering plant pigmentation and for diagnosis and treatment of prion diseases.

PS Claim 11; Page 135-136; 188pp; English.

XX

XX

CC The present sequence represents a modified N region of sup35 protein, in which the second oligopeptide repeat has been expanded twice, creating a total of seven repeats. Sup35 possesses the prion-like capacity to undergo a self-perpetuating conformational alteration that changes the functional state of Sup35 in a manner that creates a heritable change in phenotype. It is used to construct chimeric polypeptides of the invention, which

CC comprise at least one SCHAG (self-coalesces into higher-order aggregates) amino acid sequence fused in frame with a polypeptide of interest (which is other than a marker protein, a glutathione-S-transferase or a staphylococcal nuclear protein). The specification also describes chimeric polypeptides that comprises an amyloidogenic domain that causes aggregation into fibrils. The chimeric polypeptides are used to prepare polymers with multiple reactivities, e.g. derivatised with enzymes, or specific binding partners, and useful e.g. for performing multi-step chemical reactions. They can be used create an inducible, or stable phenotypic alteration in a cell, e.g. for gene therapy, protein production, imparting disease resistance to plants, altering plant dormancy and for diagnosis and treatment of prion diseases.

SQ	Sequence	271 AA:
Query	Match	97.3%; Score 682; DB 22; Length 271;
Best Local Similarity	87.2%; Pred. No. 2, 1e-62;	
Matches	123; Conservative 0; Mismatches 0; Indels 18; Gaps 1;	
Qy	1 MSDSNOGNNDQYQYSONGNQQGNNRYQOAYNAQAPAGGYYQNYCGSGYQ----- 56	
Db	1 msdsnqgnnqnyqysqsnqngnqgnrryqgyqaynaqaqpagggyqnyqggsyqggg 60	
Qy	57 -----QGYQYNPDAVYQQQNPNQGGYQQYQYQYQYQYQYQYQYQYQYQYQYQYQYQY 102	
Db	61 qgynppqgggqyqypqgggqyqypdagyyqgqyqypqgggqyqypqgggqyqypqggg 141	
Qy	103 NFNYYNNLQQGFOQFQSQG 123	
Db	121 nfnynnlnqgyqagfapqsgq 120	

ID KAB3U800 strainU; protein; 2/1 AA:
XX

RESULT 3

xxiii

AAB30/99 standard; protein; 213 A

DE A modified N region of yeast Sup35 protein

XX

KW sub3; phenotype; **SCHAG**; self-coalescence; higher-order aggregate; **amvlo**; domain; aggregation; fiber; hexagonal lattice; **alat**; **faragane**.

A modified N region of yeast Sup35 protein.

卷之二

KW: *Sup35*; phenotype; SC-HAG: self-coalesce; higher-order aggregation; amylase domain in aggregation fibril; hexamer; alloggregation.

XX

XX

XX
BD
14-DEC-2000

SACRED GAMES CENTER STAGE

PF 09-JUN-2000; 2000WO-US15876.

XX
DD
14-REFC-2000

XX

PF 09-JUN-2000; 2000WO-US15876

PIERRE BERTHIER S. J. DE MAISTRE J. B. SCHEIBELER N. SCHEIBEL T. 1822

XXI

DR WPI; 2001-061723/0

DR	WPI: 2001-061723/07.
XX	N-PSDB; AAC86684.
PT	New nucleic acid encoding chimeric proteins with self-assembly properties, useful e.g. for diagnosis and treatment of prion diseases, also related aggregates, fibrils and polymers -
PT	Example 3: Page 133-134; 188pp; English.
PS	The present sequence represents a modified N region of a yeast Sup35 protein, in which four of the five oligopeptide repeats have been deleted. Sup35 possesses the prion-like capacity to undergo a self-perpetuating conformational alteration that changes the functional state of Sup35 in a manner that creates a heritable change in phenotype. It is used to construct chimeric polypeptides of the invention which comprise at least one SC-HAG (self-coalesces into higher-order aggregates) amino acid sequence fused in frame with a polypeptide of interest (which is other than a marker protein, a glutathione-S-transferase or a staphylococcal nuclear protein). The specification also describes chimeric polypeptides that comprises an amyloidogenic domain that causes aggregation into fibrils. The chimeric polypeptides are used to prepare polymers with multiple reactivities, e.g. derivatized with enzymes, or specific binding partners, and useful, e.g. for performing multi-step chemical reactions. They can be used to create an inducible, or stable phenotypic alteration in a cell, e.g. for gene therapy, protein production, imparting disease resistance to plants, altering plant pigmentation and for diagnosis and treatment of prion diseases.
SQ	Sequence 215 AA;
Query Match	64.2%; Score 450; DB 22; Length 215;
Best Local Similarity	69.1%; Pred. No. 1.1e-38;
Matches	85; Conservative 0; Mismatches 0; Indels 38; Gaps 0
Qy	1 MSDSINOQNNOQNYQOYISQONQNOQQNRRQYQYQYNAQAPAGGYQYQGYSYQOGY
Db	1 msdsinqnqngqnyqgqsgqngqgggnnyqygqsynaqggpaggssyyqyyqyqy----- 55
Qy	61 QQYNPDAGYQQYQNPQGGYQYQOQFNPGGRGMYKNYNNNNLQGTOAGFQHQ 120
Db	56 -----Pgggrnyknfnynnnlqlqrgaqfqg 82
Qy	121 SQG 123
Db	83 sqg 85
RESULT	4
ID	AAB30820 standard; Protein; 715 AA.
XX	
AC	AAB30820;
XX	
DT	02-APR-2001 (first entry)
DE	Amino acid sequence of a <i>Candida albicans</i> protein.
XX	
KW	SCHAG; self-coalesce; higher-order aggregate; amyloidogenic domain; aggregation; fibril; phenotypic alteration; gene therapy; disease resistance; plant pigmentation; prion disease.
KW	
OS	<i>Candida albicans</i> .
PN	WO200075324-A2.
XX	
PD	14-DRG-2000.
XX	
PF	09-JUN-2000; 2000WO-US15876.
XX	
PR	09-JUN-1999; 99US-0138033.
XX	

DR PT
XX PT
XX PT
XX PT
XX PS
Claim 11; Page 173-175; 188pp; English.
The specification describes chimeric polypeptides which comprise at least one S6H6C (self-coiloses into higher-order aggregates) amino acid sequence fused in frame with a polypeptide of interest (which is other than a marker protein, a glutathione-S-transferase or a stably cooccluded nuclear protein). The specification also describes chimeric polypeptides that comprises an amyloidogenic domain that causes aggregation into fibrils. The chimeric polypeptides are used to prepare polymers with multiple reactivities, e.g., derivatised with enzymes, or specific binding partners, and useful for performing multi-step chemical reactions. They can be used to create an inducible, or stable phenotypic alteration in a cell, e.g. for gene therapy, protein production, imparting disease resistance to plants, altering plant pigmentation and for diagnosis and treatment of prion diseases. The present sequence is used in the course of the invention.
Sequence 741 AA;

ORF 1 lies between the FUS 1 or BIK 1 promoter, BIK 1 running 3' to 5'. At least one polypeptide may be promoted within a high-copy vector induced by a-factor for alpha cells, alpha-factor for a-cells using this promoter system.

XX DE
 KW DE incorporated within the HIS4 gene.
 XX
 Yeast promoter; yeast pheromones; FUS-1; BIK-1; HIS-4.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN WO8810308-A.
 XX
 PD 29-DEC-1988.
 XX
 PR 23-JUN-1988; 88WO-US02129.
 XX
 PR 24-JUN-1987; 87US-0066078.
 XX PA (WHIT-) WHITEHEAD INST.
 XX
 PI Fink GR, Trueheart J, Elion EA;
 XX
 DR WPI; 1989-023850/03.
 DR N-PSDB; AAN93100.
 XX
 PT New DNA fragment contg. protein encoding gene and yeast promoter -
 PT controlled by mating pheromone allowing efficient and
 PT regulatable expression.
 XX
 XX
 PS Claim 5; FIG 4; 51pp; English.

PR 14-MAY-1999; 99US-013370. PR 27-JUL-1999; 99US-0145918. PR 18-MAY-1999; 99US-0134768. PR 27-JUL-1999; 99US-0134941. PR 20-MAY-1999; 99US-0133124. PR 21-MAY-1999; 99US-013353. PR 24-MAY-1999; 99US-013629. PR 25-MAY-1999; 99US-0136021. PR 27-MAY-1999; 99US-013639. PR 28-MAY-1999; 99US-0136782. PR 01-JUN-1999; 99US-013622. PR 14-JUN-1999; 99US-0136119. PR 04-JUN-1999; 99US-0136502. PR 07-JUN-1999; 99US-013724. PR 08-JUN-1999; 99US-0136094. PR 10-JUN-1999; 99US-0136540. PR 10-JUN-1999; 99US-0138847. PR 14-JUN-1999; 99US-0136528. PR 16-JUN-1999; 99US-0136452. PR 16-JUN-1999; 99US-0139453. PR 17-JUN-1999; 99US-0139492. PR 18-JUN-1999; 99US-0136454. PR 18-JUN-1999; 99US-0138847. PR 18-JUN-1999; 99US-0139456. PR 18-JUN-1999; 99US-0136457. PR 18-JUN-1999; 99US-0139458. PR 18-JUN-1999; 99US-0139459. PR 18-JUN-1999; 99US-0139460. PR 18-JUN-1999; 99US-0139462. PR 18-JUN-1999; 99US-0139463. PR 18-JUN-1999; 99US-0139750. PR 18-JUN-1999; 99US-0139763. PR 21-JUN-1999; 99US-0138847. PR 22-JUN-1999; 99US-0139899. PR 23-JUN-1999; 99US-0140353. PR 23-JUN-1999; 99US-0136453. PR 24-JUN-1999; 99US-0140695. PR 28-JUN-1999; 99US-0140823. PR 29-JUN-1999; 99US-0140991. PR 30-JUN-1999; 99US-0141287. PR 01-JUL-1999; 99US-0141842. PR 01-JUL-1999; 99US-0143554. PR 02-JUL-1999; 99US-0142055. PR 06-JUL-1999; 99US-0142390. PR 08-JUL-1999; 99US-0144803. PR 09-JUL-1999; 99US-0144920. PR 12-JUL-1999; 99US-0142977. PR 13-JUL-1999; 99US-0141542. PR 14-JUL-1999; 99US-0142624. PR 15-JUL-1999; 99US-0142390. PR 16-JUL-1999; 99US-0144085. PR 16-JUL-1999; 99US-0144086. PR 19-JUL-1999; 99US-0144325. PR 19-JUL-1999; 99US-0144331. PR 20-JUL-1999; 99US-0144332. PR 20-JUL-1999; 99US-0144884. PR 21-JUL-1999; 99US-0144333. PR 21-JUL-1999; 99US-0144334. PR 19-JUL-1999; 99US-0144335. PR 20-JUL-1999; 99US-0144352. PR 20-JUL-1999; 99US-0144632. PR 20-JUL-1999; 99US-0144884. PR 21-JUL-1999; 99US-0144814. PR 21-JUL-1999; 99US-0145086. PR 21-JUL-1999; 99US-0145088. PR 22-JUL-1999; 99US-0144585. PR 22-JUL-1999; 99US-0144632. PR 22-JUL-1999; 99US-0145224. PR 23-JUL-1999; 99US-0145224. PR 26-JUL-1999; 99US-0145276.

PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145216.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134981.	PR	28-JUL-1999;	99US-0145981.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146389.
PR	24-MAY-1999;	99US-0135629.	PR	03-AUG-1999;	99US-0147038.
PR	25-MAY-1999;	99US-0136021.	PR	04-AUG-1999;	99US-0147204.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147302.
PR	28-MAY-1999;	99US-0136782.	PR	05-AUG-1999;	99US-0147192.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147260.
PR	03-JUN-1999;	99US-0137528.	PR	06-AUG-1999;	99US-0147303.
PR	04-JUN-1999;	99US-0137702.	PR	06-AUG-1999;	99US-0147416.
PR	07-JUN-1999;	99US-0137724.	PR	09-AUG-1999;	99US-0147493.
PR	08-JUN-1999;	99US-0138084.	PR	09-AUG-1999;	99US-0147795.
PR	10-JUN-1999;	99US-0138840.	PR	10-AUG-1999;	99US-0148171.
PR	10-JUN-1999;	99US-0138847.	PR	11-AUG-1999;	99US-0148319.
PR	14-JUN-1999;	99US-0139119.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139452.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139492.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139454.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139455.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139456.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139460.	PR	20-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	18-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0149930.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151303.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0151930.
PR	02-JUL-1999;	99US-0142055.	PR	07-SEP-1999;	99US-0152363.
PR	06-JUL-1999;	99US-0142390.	PR	10-SEP-1999;	99US-0153070.
PR	08-JUL-1999;	99US-0142803.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-014287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154339.
PR	13-JUL-1999;	99US-0143542.	PR	20-SEP-1999;	99US-0154779.
PR	14-JUL-1999;	99US-0143624.	PR	22-SEP-1999;	99US-0155139.
PR	15-JUL-1999;	99US-014405.	PR	23-SEP-1999;	99US-0155486.
PR	16-JUL-1999;	99US-0144085.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0144290.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	19-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	20-JUL-1999;	99US-0144332.	PR	05-OCT-1999;	99US-0157753.
PR	20-JUL-1999;	99US-0144884.	PR	06-OCT-1999;	99US-0157865.
PR	21-JUL-1999;	99US-0144884.	PR	07-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-0144333.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144334.	PR	12-OCT-1999;	99US-0158569.
PR	19-JUL-1999;	99US-0144886.	PR	13-OCT-1999;	99US-0159293.
PR	22-JUL-1999;	99US-0144335.	PR	13-OCT-1999;	99US-0159294.
PR	22-JUL-1999;	99US-0144632.	PR	18-OCT-1999;	99US-0159295.
PR	22-JUL-1999;	99US-0145089.	PR	14-OCT-1999;	99US-0159329.
PR	22-JUL-1999;	99US-0145192.	PR	14-OCT-1999;	99US-0159331.
PR	23-JUL-1999;	99US-0145145.	PR	21-OCT-1999;	99US-0160767.
PR	23-JUL-1999;	99US-0145218.	PR	21-OCT-1999;	99US-0160768.
PR			PR	21-OCT-1999;	99US-0160770.
PR			PR	21-OCT-1999;	99US-0160814.
PR			PR	21-OCT-1999;	99US-0160815.
PR			PR	22-OCT-1999;	99US-0160980.
PR			PR	22-OCT-1999;	99US-0160981.

XX
DE Tapetum-specific promoter encoded by clone pMS10.
XX
KW
XX Maize; sterility.
PN WO008831-A.

XX
PD 09-AUG-1990.
XX
PF 26-JAN-1990; 90M0-GB00114.
XX
PR 26-JAN-1989; 89GB-0001675.
XX
(ICIL) IMPERIAL CHEM IND'S PLC.

XX
PA Bridges IG, Bright SWJ, Greenland AJ, Schuch WM, Reid GA;
XX
DR N-ISDB; AAQ05759.
XX
WPI: 1990-260940/34.

PT Inhibition of gene expression in plants - which is useful in e.g.
PT imparting male sterility in maize.
XX
PS Disclosure; Fig 1; 40PP; English.

XX
CC DNA encoding the promoter is used in a gene construct with a dis-
CC rupter gene such as the mammalian uncoupling protein gene, or a
CC mutated form of the gene for the beta subunit of F1-ATPase. The
CC promoter operates in the target tissue and expression of the dis-
CC rupter gene inhibits respiration resulting in cell death. This is
CC of particular use for inhibiting male fertility in maize crops
CC where cell death results in failure to produce viable pollen. When
CC inserted into a gene cascade, as a module which permits external
CC control of expression, fertility may also be restored if required.
CC See also AAR06524 and AAR06525.
XX

Sequence 341 AA:

Query Match	19.1%	Score	134;	DB	11;	Length	341;		
Best Local Similarity	30.9%	Pred.	No. 5.9e-06;						
Matches	43;	Conservative	11;	Mismatches	55;	Indels	30;	Gaps	7;
Qy	2	SDSNQGNMQNTQYQSQNQNGNRYQGYQYINAQAPAGGYYQ-----NIQGYSY	55						
Db	126	srsrtdrnrygnqyqdgppqsgn-----yq-nnprrpppegyqgnppqgnqytysq	178						
Qy	56	QGCGYQQYNPDAGYQQ-----QYNPOCGYQ---QYNPOGGQQENQGGRCNY	101						
Db	179	qdg----rgyapqgnyaqg9qdgrgfrndytdrsqyngptdfsfqrstqyqghvnpagqqgy	236						
Qy	102	KNNYNNNL-QGQAGFQP	119						
Db	237	nnpqertufsqggggfip	255						

Search completed: February 1, 2002, 15:07:12
Job time: 317 sec

CC	-!- SUBUNIT: COMPONENT OF NUCLEAR RIBOPROTEIN COMPLEXES.	Db	132	: - : : :	159
CC	-!- SUBCELLULAR LOCATION: NUCLEAR.	FT	0-----	OPSYGQQQSYGQQSYNPPQGYGQN	
CC	-!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG ISOFORM (SHOWN HERE)	FT	0-----		
CC	AND A SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.	FT	0-----		
CC	-!- TISSUE SPECIFICITY: UBIQUITOUS.	FT	0-----		
CC	-!- DISEASE: A FORM OF MALIGNANT MYXOID LIPOSARCOMA IS CHARACTERIZED	FT	0-----		
CC	BY A CHROMOSOMAL TRANSLOCATION T(12;16) (q13;p11) THAT INVOLVES FUS	FT	0-----		
CC	AND CHP.	FT	0-----		
CC	-!- DISEASE: A FORM OF ACUTE MYELOID LEUKEMIA (AML) IS CHARACTERIZED	FT	01-JUL-1993	(Rel. 26, Created)	
CC	BY A CHROMOSOMAL TRANSLOCATION T(16;21) (p11;q22) THAT INVOLVES	FT	01-JUL-1993	(Rel. 26, Last sequence update)	
CC	FUS AND ERG.	FT	01-OCT-1996	(Rel. 34, Last annotation update)	
CC	-!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).	DE	NUCLEOPORTIN NUP100/NSP100 (NUCLEAR PORE PROTEIN NUP100/NSP100).		
CC	-!- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.	GN	NUP100 OR NSP100 OR YKL06W OR YKL336.		
CC	-!- SIMILARITY: CONTAINS 1 RAMPB-TYPE ZINC FINGER.	OS	Saccharomyces cerevisiae (Baker's yeast).		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	OC	Saccharomyctales; Saccharomycetaceae; Saccharomyces.		
CC	use by non-profit institutions as long as its content is in no way	RA	NCBI_TaxID=4932;		
CC	modified and this statement is not removed. Usage by and for commercial	RT	J. Cell Biol. 119:705-723(1992).		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/)	RN	[1]		
DR	or send an email to license@isb-sib.ch).	RP	SEQUENCE FROM N.A.		
DR	EMBL: S62138; AAB27103.1; ALT SEQ.	RX	MEDLINE=93054905; PubMed=1385442;		
DR	EMBL: X71422; CAA50558.1; ALT SEQ.	RA	Wente S.R., Rout M.P., Blobel G.;		
DR	EMBL: X71428; CAA50559.1; ALT SEQ.	RT	"A new family of yeast nuclear pore complex proteins.";		
DR	EMBL: AF071213; AAC35285.1; -.	RL	J. Cell Biol. 119:705-723(1992).		
DR	PIR: S33799; S33799.	RN	[2]		
DR	HSSP: P09551; IHAL.	RP	SEQUENCE FROM N.A.		
DR	MIM: 137070; -.	STRAIN=S288C;			
DR	MIM: 151900; -.	MBIDB=94378724; PubMed=8091863;			
DR	InterPro: IPR000504; RRM.	RASMUSSEN S.W.:			
DR	PFAM: PF00076; Rrm; 1.	RT	"Sequence of a 20.7 kb region of yeast chromosome XI includes the		
DR	SMART: SM00360; RRM; 1.	RT	nucleoside diphosphate kinase gene, tRNAs for His, Val and Trp in		
DR	SMART: SM00547; ZincF_RBZ; 1.	RT	addition to seven ORFs with weak or no significant similarity to		
DR	PROSITE: PS50102; RRM; 1.	RT	known proteins.;"		
KW	proto-oncogene; RNA-binding; DNA-binding; Nuclear protein; Repeat;	RT	Yeast 10:S59-S74(1994).		
KW	Alternative splicing; Chromosomal translocation; Zinc-finger; Zinc;	CC	-!- FUNCTION: ESSENTIAL COMPONENT OF NUCLEAR PORE COMPLEX.		
KW	Metal-binding.	CC	NUCLEOPORINS MAY BE INVOLVED IN BOTH BINDING AND TRANSLATION OF		
FT	DOMAIN 1 165 GLN/GLY/SER/TYR-RICH.	CC	THE PROTEINS DURING NUCLEOCYTOSPLASMIC TRANSPORT.		
FT	DOMAIN 166 267 GLY-RICH.	CC	-!- SUBCELLULAR LOCATION: NUCLEAR PORE COMPLEX.		
FT	DOMAIN 285 371 RNA-BINDING (RRM).	CC	-!- DOMAIN: CONTAINS G-L-F-G REPEATS.		
FT	DOMAIN 371 526 ARG/GLY-RICH.	CC	-!- SIMILARITY: BELONGS TO THE GLFG FAMILY OF NUCLEOPORINS.		
FT	ZN_FING 422 453 RANBP-TYPE.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
FT	SITE 266 267 BREAKPOINT FOR TRANSLOCATION TO FORM	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
FT	VARSPIC 64 65 PUS/TLS-CHOP ONCOGENE.	CC	use by non-profit institutions as long as its content is in no way		
FT	TG -> S (IN SHORT ISOFORM).	CC	modified and this statement is not removed. Usage by and for commercial		
SQ	SEQUENCE 526 AA: 53426 MW: 88C8E263B790549 CRC64;	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/)		
DR	DR SGD; SGD: S4518; S4518.	CC	or send an email to license@isb-sib.ch).		
DR	DR SGD; SGD: S000151; NUP100.	CC	EMBL: Z15055; CAA7875.1;		
DR	DR SGD; SGD: S000151; NUP100.	DR	EMBL: X75780; CAA53406.1; -.		
DR	DR SGD; SGD: S000151; NUP100.	DR	EMBL: Z228068; CAA81905.1; -.		
DR	DR SGD; SGD: S000151; NUP100.	DR	PIR: BA4402; BA4402.		
DR	DR SGD; SGD: S000151; NUP100.	DR	PIR: BA4402; BA4402.		
DR	DR SGD; SGD: S000151; NUP100.	DR	PIR: S339173; S339173.		
DR	DR SGD; SGD: S000151; NUP100.	DR	PIR: S339173; S339173.		
DR	DR SGD; SGD: S000151; NUP100.	DR	SGD; SGD: S000151; NUP100.		
DR	DR SGD; SGD: S000151; NUP100.	FT	Nuclear protein; Transport; Repeat.		
DR	DR SGD; SGD: S000151; NUP100.	FT	DOMAIN 33 571 29 X 6 AA APPROXIMATE REPEATS OF		
DR	DR SGD; SGD: S000151; NUP100.	FT	G-L-F-G.		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Query Match 16.0%; Score 215.5%; DB: 1; Length 959;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Best Local Similarity 28.0%; Pred. No. 5.1e-07; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	Matches 84; Conservative 37; Mismatches 110; Indels 69; Gaps 15;		
DR	DR SGD; SGD: S000151; NUP100.	FT	SEQUENCE 959 AA: 99988 MW: D398SF9901BBAAS1 CRC64;		

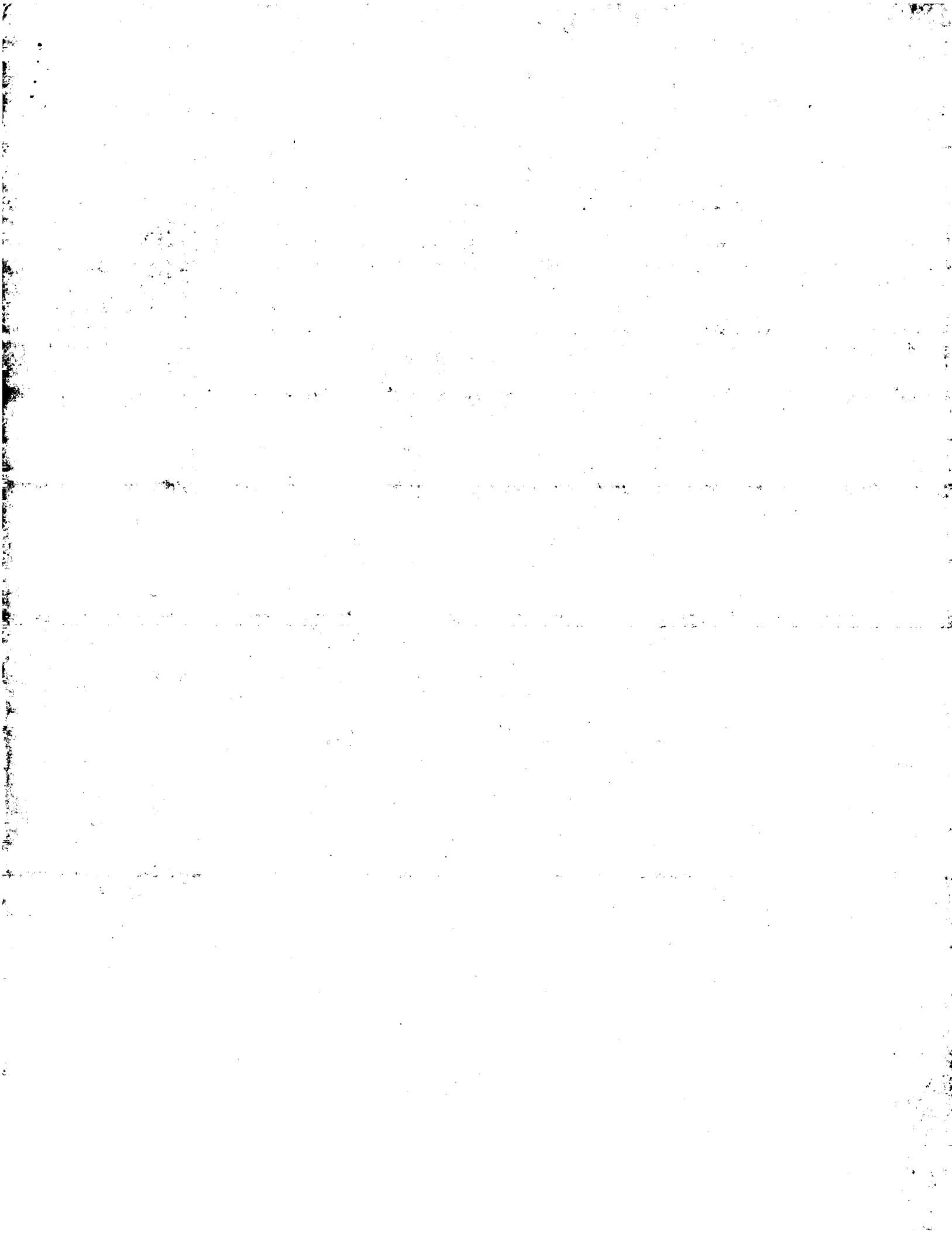
* Fri Feb 1 20:20:35 2002

us-09-591-632-50_copy_153_405.rsp

Page 11

Db 61 SGYSQSYGYENOKSSY5QPYNNQQDNMESSGSQLGRAPSIDQPY-GQDSDYDQ 119
Qy 108 OGQSGCAFSSLASMAOSYLUGGQATOSNQQOYNQQCQNNOQQYQQGQNTQHQQQQQQ 167
Db 120 SGY-----DQHGSYDEQSINYDQHDSYSQNQSYHSQRENTSHITQDRRD- 166
Qy 168 GHSSFSALASMASSYLGNNNSNSSYGQQQANYY--GRPQHNGQQSNEYGRPQYGG 224
Db 167 -----VSRYGEDNRGYGGSGGGRGGRGCGYDKDORGPMWGSSGDGGFKNFGG 214
Qy 225 NQN---SNGQHESRFNSGN 240
Db 215 HRDVGPRTDADSESPNSDNN 234

Search completed: February 1, 2002, 15:07:33
Job time: 313 sec



A; Experimental source: strain Bristol N2; clone F21C10
 Db 333 SSYLGNNNSNSNSYYGQOQANEXGRPOHNGQQQSNERYGRPQYGGNQNSNGQHESENFGSGN 392
 Qy 241 FSQONNNQNQNY 253
 Db 393 FSQONNNQNQNY 405

A; Experimental source: strain Bristol N2; clone F21C10
 C; Genetics:
 A; Gene: CESP.F21C10.8
 A; Map position: 5
 A; Introns: 20/1; 21/9/3; 24/9/3

RESULT 2
 T22330 hypothetical protein F47A4.2 - *Caenorhabditis elegans*
 C; Date: 15-oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C; Accession: T22330 submitted to the EMBL Data Library, June 1995
 A; Reference number: Z19549
 A; Accession: T22330 A; Status: preliminary; translated from GB/EMBL/DDBJ
 A; Molecule type: DNA
 A; Residues: 1-3498 <WIL>
 A; Cross-references: EMBL:Z49888; PIDN:CAA90064.1; GSPDB:GN00028; CESP:F47A4.2
 A; Experimental source: clone F47A4
 A; Introns: 36/3; 72/3; 147/2; 228/3; 261/2; 320/2; 847/3; 881/1; 1317/2; 1500/3; 1653/2;
 Query Match 17.7%; Score 239; DB 2; Length 3498;
 Best Local Similarity 31.4%; Pred. No. 1; 9e-08; DB 2; Length 3498;
 Matches 101; Conservative 20; Mismatches 117; Indels 84; Gaps 17;
 Qy 1 OSOGCQ-GOGOCGCGOGO---GOGSFATLASLASSFMNNNNNOOG----- 43
 Db 3033 QGIGSTGTGNNPGYQQQGQGQGSGSYQ-AQQQQPNQYGSNQVQGAQQQQQPLNQN 3091
 Qy 44 -NOSSGGSSFGALASMASSFMHSNNNNNNNSQGQYNSYQNCNQNSQGYNQYQGNG 102
 :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 3092 VSGSQAQAFGR-PSOSAYAVQDGGYNTG-----NOSYRDRDDOOGAODNWGSNO 3143
 Qy 103 GIVQ-----QOCQSGGAFSSLAS----MAQS--YLG-GGQTQSNOQYNOQGQ 143
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 3144 AQNQLRSQQAQAPLQQPQQSQFQQPAQAKNPMAQXQYGGFGQQGYDQQ-QQGQ 3201
 Qy 144 NNQOQYK---QOQNYOHOOQGQOQOQOCHSSSIASALAMASS-----YLGN 186
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 3202 IAPQOQAQNPQASYY-GOQQTQANQRYANGSSGYTANSGGSSNTLNQMEESGINQGSGA 3260
 Qy 187 NSNSNSYYGG---QOQNEYGRPQHNGQQNSNEYGRPQYG-----GNQNSNQHES-- 234
 :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 3261 SSNNAQGQGSSQMOQSGYGMQNMOMQMQNQKQVQRGMPCTGMQTNMGQSGMGSQMG 3320
 Qy 235 -----ENFSGNESQNNNNQ 250
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 3321 QTMQMSRSLGGGIGQQQQSQ 3342

RESULT 4
 T20847 hypothetical protein F13E9.4 - *Caenorhabditis elegans*
 C; Species: *Caenorhabditis elegans*
 C; Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C; Accession: T20847 submitted to the EMBL Data Library, February 1996
 A; Reference number: Z19332
 A; Accession: T20847 A; Status: preliminary; translated from GB/EMBL/DDBJ
 A; Molecule type: DNA
 A; Residues: 1-409 <WIL>
 A; Cross-references: EMBL:Z69383; PIDN:CAA93412.1; GSPDB:GN00022; CESP:F13E9.4
 A; Experimental source: clone F13E9
 A; Gene: CESP.F13E9.4
 A; Map position: 4
 A; Introns: 32/1; 275/3; 337/3
 Query Match 17.2%; Score 232; DB 2; Length 409;
 Best Local Similarity 27.9%; Pred. No. 6.9e-09; DB 2; Length 409;
 Matches 93; Conservative 29; Mismatches 121; Indels 90; Gaps 15;
 Qy 1 OSOGOGCQGOG---QGOG-----OGCSFTALASLASSFMNNNNNOOG---Q 43
 Db 66 QGIGQNOSSMAGSISQGQGGSMSNSPQGGYQNSQGSGMVGQYQGGS--SQSYGQAAQAOQ 122
 Qy 44 NQSGGSSFGALASMASSFMHS-----NNNNNNNSQGQYNSYQNCNQNSQGYNQYQGNG 89
 :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 123 RPPQGFQSNFGFSGQSQGSGMSNSPQGGYQNSQGSGMVGQYQGGS--SGQSGWSNLSSANS 180
 Qy 90 QGNNQHQGGNGSYQ---QOQGQGAFSSLASM-----AQSYLG 127
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 181 NGNNN---QGSSGQYNNQNGRHQGQGHHSSSSNSVMSNNGYSSNSGYNQGNNCPPTPSLN 237
 Qy 128 GGQTQSNOQY-----NQDGNNQDQYQDQGQ---QQQGQH 170
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 238 NVSSSAQDYYN1VNKKSLTTNOINEQASWASANSVQYIYETNRSAHTQATQW 297

A; Description: The sequence of C. elegans cosmid F21C10.
 A; Reference number: Z22610
 A; Accession: T22339
 A; Status: preliminary; translated from GB/EMBL/DDBJ
 A; Molecule type: DNA
 A; Residues: 1-382 <DU2>
 A; Cross-references: EMBL:U55364; PIDN:AAA97972.1; GSPDB:GN00023; CESP:F21C10.8

QY	171	SSFSALASMASSYLVGNNSNSNSSYGQQ----QANEKGRPQH-----NGQQQSNEBYGRP 220
Db	298	DINNLATVQASIEGIQNTSSSENDRQAFQALQAHPMYSVLTSLQNQEHQIYAR - 356
QY	221	OYGGNNSNSNYGNNSNSFNSVNGNNYVNHNGNNY 253
Db	357	-MGSNGNSNSNYGNNSNSFNSVNGNNYVNHNGNNY 388
RESULT	5	nuclear pore complex protein NSP116 - yeast (<i>Saccharomyces cerevisiae</i>)
N;Alternate name:	nucleoporin NSP116; protein YMRS32.12c; protein YMR047C	
C;Species:	<i>Saccharomyces cerevisiae</i>	
C;Date:	07-May-1993 #sequence_revision 07-May-1993 #text_change 21-Jul-2000	
C;Accession:	S28925; A4402; S52895; S20537	
R;Wimmer, C.; Doye, V.; Grandi, P.; Neubass, U.; Hurt, E.C.		
EMBO J. 11, 5051-5061, 1992		
A;Title:	A new subfamily of nucleoporins that functionally interact with nuclear pore proteins	
A;Reference number:	S28026; MUID:9309980	
A;Accession:	S28925	
A;Molecule type:	DNA	
A;Residues:	1-113 <WIM>	
A;Cross-references:	EMBL:X68108; NID:94052; PIDN:CAA48228.1; PID:94053	
R;Wente, S.R.; Rout, M.P.; Blobel, G.		
J. Cell Biol. 119, 705-723, 1992		
A;Title:	A new family of yeast nuclear pore complex proteins.	
A;Reference number:	A44402; MUID:93054906	
A;Accession:	A44402	
A;Molecule type:	DNA	
A;Residues:	1-25; 'A', 27-35, 'G', 537-719, 'P', 721-1017, 'Y', 1019-1022, 'Y', 1024-1113 <WEN>	
A;Note:	sequence extracted from NCBI backbone (NCBIP:117132)	
R;Odeil, C.; Bowman, S.		
A;Reference number:	S52885	
A;Accession:	S52896	
A;Molecule type:	DNA	
A;Residues:	1-113 <CODE>	
A;Cross-references:	EMBL:Z48502; NID:9695715; PID:9695727; MIPS:YMR047C	
A;Genes:	SGD:NUP116; NSP116	
A;Cross-references:	SGD:S0004650; MIPS:YMR047C	
C;Keywords:	nucleus	
Query Match	16.7%	Score 224.5; DB 2; Length 1113;
Best Local Similarity	29.1%	Pred. No. 5.8e-08;
Matches	95;	Conservative 18; Mismatches 17; Indels 117; Gaps 16;
QY	6	GOGQOGQOGQOGQOGQGFTAAKLASSFMNSNNNOQGQNSGGSSPGALAMASSFMHS 65
Db	217	GTGTGGGGGGGGATNSTGL-----FGSSTN-----LSGNNSARGANKPATSGCLF 262
QY	66	NNNQNSNSNSO-----GYNOXYONG-----NONSOYNN-----QOYQ 98
Db	263	NTNNNPNTGNTNTGLFGQQNSNTNGGLFGQQNSFGANNVNGGAQFQVNIRGAPOQQTQ 322
QY	99	GGNGG-----YQQQGQ-----QSGQAFSSLA-----SMAQYLG 128
Db	323	OGGGIFGQSNANANGGAFCQOOQGTGALFGAKPASGGLFQ30SAGSKAFGMNTNPQTG 382
QY	129	GQTQSNQQ-----YQQQ-----GQNNQQQKQ-----GONYQHQQQGQQQGHSS 172
Db	383	LFGOTNQOQGGGLFGQQQNSNAGGLFGQONNQNSQSGLFGQONNSNAFGPQOQGLFG 442
RESULT	7	glutenin high molecular weight chain 1By9 precursor - wheat
C;Species:	<i>Triticum aestivum</i> (common wheat)	
C;Date:	08-Jun-1994 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999	
C;Accession:	S18733	
R;Halford, N.G.; Forde, J.J.; Anderson, O.D.; Greene, F.C.; Shawty, P.R.		
Theor. Appl. Genet. 75, 11-126, 1987		
A;Title:	The nucleotide and deduced amino acid sequences of an HMW glutenin subunit g	
A;Reference number:	S18733	
A;Accession:	S18733	
A;Status:	preliminary	
A;Molecule type:	DNA	
A;Residues:	1-705 <HAL>	
A;Cross-references:	EMBL:X61026; NID:922089; PIDN:CAA43361.1; PID:922090	
C;Superfamily:	glutenin	
Db	443	SKPAGGLFGQQGASSFASQNAQNSIFG-----QNNQQQOST-----GGLF 484
QY	224	GNQNSQHQHSEFNESNSFSQONNINGO 250

F;22-838/Product: glutenin, HMW chain #status predicted <MAT>

Query Match 16.1%; Score 216.5; DB 1; Length 838;
 Best Local Similarity 29.9%; Pred. No. 1.5e-07;
 Matches 90; Conservative 14; Mismatches 132; Indels 65; Gaps 12

Db 189 QGQPQGQQ-QGRQPQGQGPQGYPTSSLQQLPQGLQQPAQOGQOGQOPGQQQG----- 239

Db	34	GQOSY---SOSQSRTDTSYGQGQSSYSSGQSNTGTYGTQSPTPGIGSTGGVGQSSQQS	89
Qy	153	GQNYHQHQPSQQQQPQGHSSFSALASMASSYLGNNNSNNSYGGQDANEYGRPQHNGQQ	212
Db	90	SYGQQSSYPGYQGPAPS-----TSGSYGSS50SSSYGQPQSGSYS	131
Qy	213	QSNEYGRPQIGGNQNNSNGQHESFNFGNSGNFQON	245
Db	132	Q----QPS1GGQQSYGQQSYNPNPQGYGQON	159

Qy	60	SSFMHSNNNQNNSNNSDQGYNQNSYQNQNONNSQGYNQNQOYQGGNGGY-----QQQGQSGG	11
Eb	240	--QQPCQGQQPGQGQGQG--QQPGQGQQPGQGQGQGQLQQGQQCYYPTSLQQSGQGPY	29
Cy	114	AFSSLASMAQSVLG-----GGQTQSNNQQYNNQQGSNNQ---QQYQQGQNYHQHQGQ	16

RESULT 15
S33799
RNA-binding protein FUS, nuclear - human
N; Alternate names: RNA-binding protein TLS
C; Species: Homo sapiens (man)

Db	295	YPTSTLQDQGQSGSYVPTSPQQPGQGQPGQQLQPGQPAQGQPEAQGQGQPGQGQGQGP
Cy	163	QQQQGHSSFSALASMASASYLG-----NNNSNSNSYYGG-----QANAEVG 204
Db	355	QQQQPGQGPQGYPTSPQQSGCQGPQGYPTSSQQOPTSQQQPGQGQGQGQVQGQQAQEG 414
Cy	205	RPOHNGQQQSNEY-----GRPOY---GENONISNGOHE--SENNSIGNEQQNNING 249
Db	415	QQQQPGQGPQGYPTSPQQSGCQGPQGYPTSSQQOPTSQQQPGQGQGQGQVQGQQAQEG 474
Cy	250	-
Db	475	0 475

C;Accession: S33793; S36157
Nature 333, 640-641, 1993
A;Title: Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma.
A;Reference number: S33798; MUID:93288139
A;Accession: S33799
A;Molecule type: mRNA
A;Residues: 1-526 <CRO>
A;Cross-references: GB:S62140; NID:9386156; PIDN:AAB27102.1; PID:9386157
A;Experimental source: liposarcoma
R;Rabbits, T.H.; Forster, A.; Larson, R.; Nathan, P.
Nature Genet. 4, 175-180, 1993
A;Title: Fusion of the dominant negative transcription regulator CHOP with a novel gene product, p55, creates a dominant-negative protein, p55^{CHOP}, that inhibits the p55^{CHOP} protein.

RESULT 14
S33798
FUS/CHOP mutant fusion protein - human

A; Status: preliminary
A; Molecule type: mRNA
A; Residues: 1-63 , 'S' 66-526 <RAB>

```
C;Species: Homo sapiens (man)
C;Sequence: S13798: S3615A
C;Date: 22-Nov-1993 #sequence_revision 30-Nov-1995 #text_change 08-Dec-2000
C;Accession: S13798
```

A;Experimental source: liposarcoma
A;Genetics:
A;Gene: CNA1Q1

Nature 363 640-644, 1993
H;Title: Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma.
H;Reference number: S33798; MUID:93288139
H;Accession: S33798
H;Molecule type: mRNA
H;Residues: 1-462 <CRO>
H;Cross-references: GR:S62138; NID:0386158; RTDN:AA827103; 1; RTD:0386150

A;MacCoss:references: GDB1136048; OMIM:137070
C;Position: 1bp11.2-16p11.2
C;Function:
A;Description: RNA binding; probable plays a role in transcriptional regulation
C;Superfamily: RNA-binding protein EWS; ribonucleoprotein repeat homology
C;Keywords: carcinogenesis; nucleus; RNA binding; tandem repeat
F;286-361/Domain: ribonucleoprotein repeat homology <RRM>

Nature Genet. 4, 177-180, 1993
A;Title: Fusion of the dominant negative transcription regulator CHOP with a novel gene
A;Reference number: S36157
A;Accession: S36158
A;Molecule type: mRNA
A;Residues: 1-662 <RAB>
A;Cross-references: EMBL:X71427; NID:9395919; PIDN:CA50558.1; PID:94210364
C;Comment: This sequence is the chimeric product of a translocation mutation.
C;Genetics:
A;Gene: GADB153:FUS
A;Map position: 12q33/16p11
A;Note: TLS is a synonym for GDB:FUS
C;Keywords: fusion protein

Query Match 16.0%; Score 216; DB 4; Length 462;
 Best Local Similarity 30.5%; Pred. No. 9 2e-08;
 Matches 65; Conservative 15; Mismatches 79; Indels 54; Gaps 6;

Search completed: February 1, 2002, 15:02:49
Job time: 219 sec

Search completed: February 1, 2002, 15:02:49

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 1, 2002, 15:02:16 ; Search time 19.74 Seconds

(without alignments)
288.416 Million cell updates/sec

Title: US-09-591-632-50_COPY_153_405
Perfect score: 1347
Sequence: 1 QGQGQQGQGOGQGQGCGQG... SFNFSGNFSQQNNNGNQRY 253
Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503592 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/iaa/5A_COMBO.pep: *
 2: /cgn2_6/ptodata/2/iaa/5B_COMBO.pep: *
 3: /cgn2_6/ptodata/2/iaa/6A_COMBO.pep: *
 4: /cgn2_6/ptodata/2/iaa/6B_COMBO.pep: *
 5: /cgn2_6/ptodata/2/iaa/PCUS_COMBO.pep: *
 6: /cgn2_6/ptodata/2/iaa/backtitles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

RESULT		ALIGNMENTS	
1	PCP-US93-03027-3	Sequence 3, Application PC/TUS9303027	
		GENERAL INFORMATION:	
		APPLICANT: LEONARD, WARREN; TOLEDANO,	
		APPLICANT: MICHEL	
		TITLE OF INVENTION: PREVENTION OF BINDING OF NF-B/REL/DORSAL	
		TITLE OF INVENTION:	
		NUMBER OF INVENTIONS: 9	
		CORRESPONDENCE ADDRESS:	
		ADDRESSEE: MORGAN & FINNEGAN	
		STREET: 345 PARK AVENUE	
		CITY: NEW YORK	
		STATE: NEW YORK	
		COUNTY: USA	
		ZIP: 10154	
		COMPUTER READABLE FORM:	
		MEDIUM TYPE: FLOPPY DISK	
		COMPUTER: IBM PC COMPATIBLE	
		OPERATING SYSTEM: PC-DOS/MS-DOS	
		SOFTWARE: WORDPERFECT 5.1	
		CURRENT APPLICATION DATA:	
		APPLICATION NUMBER: PCT/US93/03027	
		FILING DATE: 19930401	
		PRIOR APPLICATION DATA:	
		APPLICATION NUMBER: US/07/862,987	
		FILING DATE: 06-APR-1992	
		ATTORNEY/AGENT INFORMATION:	
		NAME: DOROTHY R. AUTH	
		REGISTRATION NUMBER: P-36,434	
		REFERENCE/DOCKET NUMBER: 2026-4010 PCT	
		TELECOMMUNICATION INFORMATION:	
		TELEPHONE: 212-758-4800	
		TELEFAX: 212-751-6849	
		INFORMATION FOR SEQ ID NO: 3:	
		SEQUENCE CHARACTERISTICS:	
		LENGTH: 678	
		TYPE: AMINO ACID	
		STRANDEDNESS: single	
		TOPOLOGY: linear	
		MOLECULE TYPE: protein	
		HYPOTHETICAL: No	
		ORIGINAL SOURCE:	
		ORGANISM: Drosophila melanogaster	
		STRAIN: Oregon R	
		INDIVIDUAL ISOLATE:	
		DEVELOPMENTAL STAGE: embryo	
		HAPLOTYPE:	
		TISSUE TYPE:	

1 129

2

US-08-317-848B-2

29

747

3

US-09-034-177-3

30

126.5

9.4

341

2

US-08-538-711A-8

31

126.5

9.4

341

4

US-08-735-027-8

32

126.5

9.4

353

2

US-08-538-711A-7

33

126.5

9.4

353

4

US-08-725-021-7

34

126.5

9.4

414

1

US-08-313-682-1

35

126.5

9.4

414

2

US-08-705-660-26

36

126.5

9.4

591

3

US-08-935-903B-2

37

126

9.4

428

1

US-08-190-02A-29

38

125.5

9.3

428

4

US-08-417-346-29

39

125.5

9.3

428

4

US-08-343-682-1

40

125

9.3

656

4

US-09-244-56A-4

41

9.3

656

1

US-08-414-139-9

42

125

9.3

1250

1

US-08-414-139-9

43

123

9.1

606

4

US-08-536-978B-21

44

123

9.1

606

4

US-09-247-806-4

45

122.5

9.1

353

2

US-08-681-702-37

36

122.5

9.1

353

2

US-08-317-848B-2

37

122.5

9.1

353

2

US-08-425-069-2

38

122.5

9.1

353

2

US-08-425-069-2

39

122.5

9.1

353

2

US-08-425-069-2

40

122.5

9.1

353

2

US-08-425-069-2

41

122.5

9.1

353

2

US-08-425-069-2

42

122.5

9.1

353

2

US-08-425-069-2

43

122.5

9.1

353

2

US-08-425-069-2

44

122.5

9.1

353

2

US-08-425-069-2

45

122.5

9.1

353

2

US-08-425-069-2

46

122.5

9.1

353

2

US-08-425-069-2

47

122.5

9.1

353

2

US-08-425-069-2

48

122.5

9.1

353

2

US-08-425-069-2

49

122.5

9.1

353

2

US-08-425-069-2

50

122.5

9.1

353

2

US-08-425-069-2

51

122.5

9.1

353

2

US-08-425-069-2

52

122.5

9.1

353

2

US-08-425-069-2

53

122.5

9.1

353

2

US-08-425-069-2

54

122.5

9.1

353

2

US-08-425-069-2

55

122.5

9.1

353

2

US-08-425-069-2

56

122.5

9.1

353

2

US-08-425-069-2

57

122.5

9.1

353

2

US-08-425-069-2

58

122.5

9.1

353

2

US-08-425-069-2

59

122.5

9.1

353

2

US-08-425-069-2

60

122.5

9.1

353

2

US-08-425-069-2

61

122.5

9.1

353

2

US-08-425-069-2

TELECOMMUNICATION INFORMATION: 2

TELEPHONE: (510) 923-3888
TELEFAX: (510) 655-3442
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 903 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-853-310-2

Query Match 11.8%; Score 158.5; DB 2; Length 903;
Best Local Similarity 28.7%; Pred. No. 2.8e-06; Matches 81; Conservative 21; Mismatches 121; Indels 59; Gaps 10;

Qy 5 QGGCGQGCGQGCGQGCGSFTALASSFMNSNNNNQQQ - NQSSGGSSFGALASMA 59
Db 97 QOHQQQONQOQHQLHDYDVOVLAQHSPMPOAHSEFGNDIGQEMILCDAPVPMASAEMEV 156

Qy 61 SFM--HSNNNQNNSN----SOQGYNOSYONGNQNSQGYNNQYQ---- 98
Db 157 STIVTNSSNSNDSSNNISLCSSTNSLTINMPHQASQQFOQNAQSNAAQORQILVDSNGQ 216

Qy 99 -GGNGGYDQOGQGCGGAFSSLASMAQSYLGGGOTOSNQDQYQDQGNNQOQYQDQGNYQ 157
Db 217 IIGNFLQOORQO----- 254

Qy 158 HQQCGQQQOQGHSSF----SALASMASYLVGNNSNSNNSYQGQQQANAEYGRPHNSQ 211
Db 255 HQQQQQQQQATSSNSLGLKTPVALRNQGTOFLSPNLQHQHQQQQQLEQHQQQATAQ 314

Qy 212 QQSNEYGRPGQYGGNONSNGHE--SFNFSGNFSQSNNNNGQN 251
Db 315 QKHQOI--QPFALAOQALHQRLAQAANINLQQQQQQQN 354

RESULT 6
US-08-219-849-34
; Sequence 34, Application US/09219849
; Patent No. 6150091
; GENERAL INFORMATION:
; APPLICANT: VAN HEERDE, GEORGE V.
; APPLICANT: VAN RIJN, ALEXIS C.
; APPLICANT: BOUNSTRA, JAN B.
; APPLICANT: DE WOLF, FREDERIK A.
; APPLICANT: MOOROOK, ANDREAS
; APPLICANT: WERPEN, MARC W.T.
; APPLICANT: WIND, RICHIE D.
; TITLE OF INVENTION: SILVER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN
; TITLE OF INVENTION: SUITABLE FOR PHOTOGRAPHIC APPLICATION AND ALSO THE
; TITLE OF INVENTION: PREPARATION THEREOF
; FILE REFERENCE: 2728-2
; CURRENT APPLICATION NUMBER: US/09/247,806
; NUMBER OF SEQ ID NOS: 50
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Nepilia clavipes
; US-09-247-806-2

Query Match 11.5%; Score 155.5; DB 4; Length 529;
Best Local Similarity 27.7%; Pred. No. 2.7e-06; Matches 75; Conservative 12; Mismatches 105; Indels 79; Gaps 42;

Qy 2 GQGCGQ---GQGCGQGCGQGCGSFTALASSFMNSNNNNQQG - 42
Db 3 GYGPQGPGCGYGPQGPGCGPSPGSPGSSAAAAAAAAGFGGYGPQGQGPGGYGPQGQGPGRYG 62

Qy 43 --QNOQSGGSSFGALASMASSFMHSNNNNNSNSQGYQYQSYONGNQNSQGYNNQYQG 99
Db 63 PGQQPSGPSSPSAAMAAAG----- 94

Qy 100 GNGGYQDQGCGGAFSSLASMAQSYLGGGOTOSNQDQYQDQGNNQOQYQDQGNYQ 158
Db 95 GPGGGQGQGCGPSPGCSAAMASASAESGGPGGYGP---GQGPGGYGPQGQGPGGY-- 149

Qy 159 QQGQQQQQGHSSFSALASMASSYLVGNNSNSNNSYQ-GQOQANAEYGRPHN---- 209
Db 150 -GPGGGQGPGSPGCSAAMASASAESGGPGGYGPQGQGPGGYGP3QGPSPGPGSA 204

Qy 210 -----GQOSNEYGRPGQYGGNONSNGQ 231

Sequence 2, Application US/08749391
Patent No. 5948667
GENERAL INFORMATION.

GENERAL INFORMATION:
APPLICANT: Cheng, Kuo-Joan
APP'TANT: Selinger, Leonard B

APPLICANT: Hu, Youji
APPLICANT: Forsberg, Cecil W.
APPLICANT: Moloney, Maurice M.
TITLE OF INVENTION: A Xylanase Obtained From an
TITLE OF INVENTION: Anaerobic Fungus
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:

APPLICANT: Moloney, Maurice M.
TITLE OF INVENTION: A Xyloanase Obtained From an
TITLE OF INVENTION: Anaerobic Fungus
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greene, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201

ZIP: 80803
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US08/749,391
FILING DATE: 13-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Donna M. Ferber
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 93-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 485 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

1. MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOSS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.3.C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/390,200
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/749,391
FILING DATE: 13-NOV-1996
ATTORNEY / AGENT INFORMATION:
NAME: Donna M. Ferrier
REGISTRATION NUMBER: 33,878
REFERENCE / DOCKET NUMBER: 93-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 485 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-02-390-200-2

RESULT 9
US-09-390-200-2
; Sequence 2, Application US/09390200
; Patent No. 6137032

RESULT 12
US-08-328-809-8
; Sequence 8, application US/08328009
; Patent No. 5705334

GENERAL INFORMATION:

APPLICANT: Lippard, Stephen J.
APPLICANT: Essigmann, John M.
APPLICANT: Donahue, Brian A.
APPLICANT: Toney, Jeffrey H.
APPLICANT: Bruun, Suzanne L.
APPLICANT: Pil, Pieter M.
APPLICANT: Brown, Steven
APPLICANT: Kelllett, Patti

TITLE OF INVENTION: Uses For DNA Structure-Specific
Title of Invention: Recognition Proteins

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patent Administrator, Testa, Hurwitz & Thibeault
STREET: 53 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/328,809
FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fenton, Gillian M.
REGISTRATION NUMBER: 36,508
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-248-7100
TELEFAX: 617-248-7100

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 542 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide

ORIGINAL SOURCE:
ORGANISM: *Saccharomyces cerevisiae*
IMMEDIATE SOURCE:
CLONE: fractional yeast SSRP (fySSRP) (predicted)

US-08-328-809-8

Query Match 10.8%; Score 145; DB 1; Length 542;
Best Local Similarity 26.2%; Pred. No. 2.4e-05;
Matches 68; Conservative 15; Mismatches 81; Indels 96; Gaps 9;

Oy 1 OCGOGOGOGOGOGOGOGOGSP-----TALASAS-----SFMSN 36
Oy 9 QQQQQQQQQQQQQQQQPYQHQQSPQQQQVYFPLPPQSLQTSQSQQQQQYANSN 68
Db 37 NNNQCONQSSGSGSICALASHAS-----SFM----- 63

RESULT 13
PCT-US92-11107-13
; Sequence 13, Application PCT/US92/11107

GENERAL INFORMATION:

APPLICANT: Donahue, Brian A.
APPLICANT: Toney, Jeffrey H.
APPLICANT: Bruun, Suzanne L.
APPLICANT: Pil, Pieter M.
APPLICANT: Brown, Steven
APPLICANT: Kelllett, Patti
APPLICANT: Essigmann, John M.

TITLE OF INVENTION: DNA Structure Specific Recognition
Title of Invention: Protein and Uses Therefor

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: 2 Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/11107
FILING DATE: 1992/12/18

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/539, 906
FILING DATE: 18-JUN-1990

ATTORNEY/AGENT INFORMATION:

NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-4787AAA

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 542 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide

ORIGINAL SOURCE:
ORGANISM: *Saccharomyces cerevisiae*
IMMEDIATE SOURCE:
CLONE: fractional yeast SSRP (fySSRP) (predicted)

PCT-US92-11107-13

Db 69 SNSNNNNVNVALPQDFGYMQGTGSGONYPTINODQFSEFYNSLFLHTOKTNPSVTGIG 128
Qy 64 -HSNNNQNSNNNSQGQNYQSYNGQ-----NSQGNNNOQYQGGNGYQOOQGQ- 110
Db 129 ASSNNNSNNNVSSGNSTSSNPTQLAASQINPATATAANN--AAGPASYLSQLPQV 185
Qy 111 -----SGGASSL--ASMAQSYLGGQTS-----NQQYNQGQNNQQ 147
Db 186 QRYVYNMNAALSSLDPSSAGNAAGNANTATHPGLLPPNLQPQLTHHQOQMQQQLQLQQ 245
Qy 148 QYQQGQNYHQDQQQGQQQ 167
Db 246 QQLOQQQQQHQDQQQ 265
Qy 148 QYQQGQNYHQDQQQGQQQ 167
Db 246 QOLOQQQQQHQDQQQ 265

Db 69 SNSNNNNVNVALPQDFGYMQGTGSGONYPTINODQFSEFYNSLFLHTOKTNPSVTGIG 128
Qy 64 -HSNNNQNSNNNSQGQNYQSYNGQ-----NSQGNNNOQYQGGNGYQOOQGQ- 110
Db 129 ASSNNNSNNNVSSGNSTSSNPTQLAASQINPATATAANN--AAGPASYLSQLPQV 185
Qy 111 -----SGGASSL--ASMAQSYLGGQTS-----NQQYNQGQNNQQ 147
Db 186 QRYVYNMNAALSSLDPSSAGNAAGNANTATHPGLLPPNLQPQLTHHQOQMQQQLQLQQ 245
Qy 148 QYQQGQNYHQDQQQGQQQ 167

Db 69 SNSNNNNVNVALPQDFGYMQGTGSGONYPTINODQFSEFYNSLFLHTOKTNPSVTGIG 128
Qy 64 -HSNNNQNSNNNSQGQNYQSYNGQ-----NSQGNNNOQYQGGNGYQOOQGQ- 110
Db 129 ASSNNNSNNNVSSGNSTSSNPTQLAASQINPATATAANN--AAGPASYLSQLPQV 185
Qy 111 -----SGGASSL--ASMAQSYLGGQTS-----NQQYNQGQNNQQ 147
Db 186 QRYVYNMNAALSSLDPSSAGNAAGNANTATHPGLLPPNLQPQLTHHQOQMQQQLQLQQ 245
Qy 148 QYQQGQNYHQDQQQGQQQ 167

Db 69 SNSNNNNVNVALPQDFGYMQGTGSGONYPTINODQFSEFYNSLFLHTOKTNPSVTGIG 128
Qy 64 -HSNNNQNSNNNSQGQNYQSYNGQ-----NSQGNNNOQYQGGNGYQOOQGQ- 110
Db 129 ASSNNNSNNNVSSGNSTSSNPTQLAASQINPATATAANN--AAGPASYLSQLPQV 185
Qy 111 -----SGGASSL--ASMAQSYLGGQTS-----NQQYNQGQNNQQ 147
Db 186 QRYVYNMNAALSSLDPSSAGNAAGNANTATHPGLLPPNLQPQLTHHQOQMQQQLQLQQ 245
Qy 148 QYQQGQNYHQDQQQGQQQ 167

Db 69 SNSNNNNVNVALPQDFGYMQGTGSGONYPTINODQFSEFYNSLFLHTOKTNPSVTGIG 128
Qy 64 -HSNNNQNSNNNSQGQNYQSYNGQ-----NSQGNNNOQYQGGNGYQOOQGQ- 110
Db 129 ASSNNNSNNNVSSGNSTSSNPTQLAASQINPATATAANN--AAGPASYLSQLPQV 185
Qy 111 -----SGGASSL--ASMAQSYLGGQTS-----NQQYNQGQNNQQ 147
Db 186 QRYVYNMNAALSSLDPSSAGNAAGNANTATHPGLLPPNLQPQLTHHQOQMQQQLQLQQ 245
Qy 148 QYQQGQNYHQDQQQGQQQ 167

Db 69 SNSNNNNVNVALPQDFGYMQGTGSGONYPTINODQFSEFYNSLFLHTOKTNPSVTGIG 128
Qy 64 -HSNNNQNSNNNSQGQNYQSYNGQ-----NSQGNNNOQYQGGNGYQOOQGQ- 110
Db 129 ASSNNNSNNNVSSGNSTSSNPTQLAASQINPATATAANN--AAGPASYLSQLPQV 185
Qy 111 -----SGGASSL--ASMAQSYLGGQTS-----NQQYNQGQNNQQ 147
Db 186 QRYVYNMNAALSSLDPSSAGNAAGNANTATHPGLLPPNLQPQLTHHQOQMQQQLQLQQ 245
Qy 148 QYQQGQNYHQDQQQGQQQ 167

Db 69 SNSNNNNVNVALPQDFGYMQGTGSGONYPTINODQFSEFYNSLFLHTOKTNPSVTGIG 128
Qy 64 -HSNNNQNSNNNSQGQNYQSYNGQ-----NSQGNNNOQYQGGNGYQOOQGQ- 110
Db 129 ASSNNNSNNNVSSGNSTSSNPTQLAASQINPATATAANN--AAGPASYLSQLPQV 185
Qy 111 -----SGGASSL--ASMAQSYLGGQTS-----NQQYNQGQNNQQ 147
Db 186 QRYVYNMNAALSSLDPSSAGNAAGNANTATHPGLLPPNLQPQLTHHQOQMQQQLQLQQ 245
Qy 148 QYQQGQNYHQDQQQGQQQ 167

Db 69 SNSNNNNVNVALPQDFGYMQGTGSGONYPTINODQFSEFYNSLFLHTOKTNPSVTGIG 128
Qy 64 -HSNNNQNSNNNSQGQNYQSYNGQ-----NSQGNNNOQYQGGNGYQOOQGQ- 110
Db 129 ASSNNNSNNNVSSGNSTSSNPTQLAASQINPATATAANN--AAGPASYLSQLPQV 185
Qy 111 -----SGGASSL--ASMAQSYLGGQTS-----NQQYNQGQNNQQ 147
Db 186 QRYVYNMNAALSSLDPSSAGNAAGNANTATHPGLLPPNLQPQLTHHQOQMQQQLQLQQ 245
Qy 148 QYQQGQNYHQDQQQGQQQ 167

FT Misc-difference 40
 FT /note= "microsequenced and derived SEQ identical"
 XX WO9515377-A.
 PN XX 08-JUN-1995.
 PD XX 01-DEC-1994; 94WO-AU00744.
 PR XX 18-JUL-1994; 94US-0276452.
 PR 03-DEC-1993; 93US-0161944.
 PA XX (ALBR) ALBRIGHT & WILSON AUSTRALIA.
 PA Bacic A, Chen C, Clarke AB, Du H, Gane AM, Mau S;
 XX DR WPI; 1995-212528/28.
 DR N_PSDB; AAQ25254.
 XX PN Recombinant plant arabinogalactan protein and cloned DNA encoding it -
 PT useful as an emulsifying agent, adhesive agent or a lubricating agent.
 XX 2S
Claim 26; Figure 1F; 142pp; English.
 XX The N-terminal region of an arabinogalactan protein (AGP) was obt'd. from *N*. *alata* suspension culture. It was used as a template for the synthesis of an oligo primer which was further utilised for the isolation of a hybridising AGP gene from both *N. alata* and *N. plumbea* *Inafolia*. Degenerate reverse primers corresp. to part of the AGP N-terminal AA sequence were synthesised and used to obtain a 160 bp primer extension product, which was then amplified by PCR. Additional primers, corres. in sequence to parts of the 160 bp fragment, were synthesised and used to amplify the 3' part of the AGP gene by nested PCR. The primers were AAQ92536 and AAQ92537. A 1.6 kb fragment was amplified and sequenced. The alignment of the sequences obt'd. from the two PCR reactions gave rise to a DNA sequence shown. The PCR sequencing with two mismatches (see AAQ92524 FT). The 1.6 kb PCR fragment was used to screen a cDNA library made from RNA isolated from *N. alata* cells. The alignment of the PCR sequences w/ the cDNA sequence gave rise to AAQ92524. The calc'd. mol. wt. of the encoded protein is 51.8 kD with a predicted pI of 3.84.
 CC SQ Sequence 461 AA;

Query Match 15.1%; Score 203.5; DB 16; Length 461; Best Local Similarity 27.5%; Pred. No. 2.8e-09; Matches 66; Conservative 33; Mismatches 96; Indels 45; Gaps 13;

Qy 34 NSNNNNQOG--ONQSSGSSRGALASMASSEFHNSHNNSN---NSQGQNQSYQDNGQ 87
 Db 195 nnnnnndgdrssyenyenankunngyseynnnnnnngyaknnyngysqsynnnn 254
 Qy 88 -NSQGYNNQ---QYQGGNGYQQ---QGQSGGAFLSLASMAQSYLGGOTQSN--- 134
 Db 255 fysenyennnnnvnsensnnngysskkinngyssqyymnnngfssyynnnnnnnnnv 314
 Qy 135 QQYQNGQGQNN -QQQQQQGQNYHQHQQGQQQGQHSSSALSASASSVIGNNNSNS 192
 Db 315 snyenynnnnnvseynnnnnnafyenyunnngysseny---nqassy---nnndt 367
 Qy 193 SYGGQANEXGRQHNGQQ---SNEYGRPOYGGNNSQHESENFSFGNSFQSQNNNG 249
 Db 368 v---erqglsdtrflengkyydkine-----ntnnnyseny---hvssyunn 413

XX 26-MAR-2001 (first entry)
 DT DE Gene 21 human secreted protein homologous amino acid sequence #123.
 XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiot; vasotropi; cerebroprotective; nootropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm; rheumatoid arthritis; hyperproliferative disorder; cardiac arrest; cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia; angiogenesis; nervous system disorder; Alzheimer's disease; infection; ocular disorder; corneal infection; wound healing; skin aging; food additive; preservative.
 KW Ruminococcus flavifaciens.
 OS PN WO20061629-A1.
 XX DR 19-OCT-2000.
 PD XX 06-APR-2000; 2000WO-US09071.
 PR XX 09-APR-1999; 99US-0128694.
 PR XX 20-JAN-2000; 2000US-0176931.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX PI Ruben SM, Komatsoulis G;
 XX DR WPI; 2000-647420/62.
 PT XX Isolated nucleic acid molecule encoding a human secreted protein is used in preventing, treating or ameliorating a medical condition -
 PS Disclosure; page 500-501; 533pp; English.
 CC AAFT2373 to AF22421 encode the human secreted proteins given in AAB63134 to AAB63182. AAB63183 to AAB63231 represent more human secreted proteins and polypeptides homologous to them. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Examples of activities include: immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; anticancer; vasotropi; cerebroprotective; nootropic; neuroprotective; antibacterial; virucide; fungicide; and ophthalmological. The polynucleotides and proteins can be used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. AF22364 to AF22372 and AAB63133 represent sequences used in the exemplification of the present invention.

XX SQ Sequence 280 AA;

Query Match 14.7%; Score 198.5; DB 21; Length 280; Best Local Similarity 26.7%; Pred. No. 4.1e-09; Matches 69; Conservative 32; Mismatches 104; Indels 53; Gaps 9;

Qy 7 QSGQGQGQGQGQGQSFALLASLASSFMNSNNNQGQNSQSGSSFCALLASHASSFHNSN 66
 AC AAB63197;

ID AAB69614 standard; Protein; 145 AA.
NC
XX
AAB69614;
XX

RESULT 10
 AAB35148
 ID AAB35148 standard; Protein; 894 AA.
 XX
 AC AAB35148;
 XX
 DT 09-APR-2001 (first entry)

KW Neurological disorder; Huntington's disease; Alzheimer's disease; Parkinson's disease; prion disease; frontotemporal dementia; amyotrophic lateral sclerosis; spinal and bulbar muscular atrophy;

DE Human NFAR-2 SEQ ID NO: 4.
 XX Human; nuclear factor associated with dsRNA; NFAR-1; NFAR-2;
 KW transcription regulator; chromosome 19p13.1-13.2; apoptosis;
 KW tumourigenesis.
 XX
 OS Homo sapiens.

PN
XX
PD
XX
PF
WO200106989-A2.
01-FEB-2001.
24-JUL-2000; 2000WO-US20131.

PN
XX
PD
XX
PF
XX
PR
XX
WO2000077205-ALL.
21-DEC-2000
09-JUN-2000; 2000WO-US15767
11-JUN-1999; 99US-0138612

XX
 ?A (HUST/) HUSTON J S.
 PA (MESS/) MESSE A.
 PA (LECE/) LECERF J.

PA (BARB,) BARBER G N.
PA (SAUN,) SAUNDERS L.
PA (PERK,) PERKINS D J.
XX
PT Barber GN, Saunders L, Perkins DJ;
XX DR WPI: 2001-080888/09.
PT " DPO 2001-080888/09.

XX
JK
WPI, 2001-162/007/16.
PT
Inhibiting intracellular polypeptide accumulation, useful for treating
neurological disorders, e.g., Alzheimer's disease, comprises contacting
the nervous tissue with a monoclonal antibody.

XX
PT Novel isolated human nuclear factor associated with dsRNA polypeptide
PT useful for determining structure-function relationships and as affinity
PT tag to identify and isolate interacting proteins that bind to the
PR factor -
XX
PS Claim 2: page 50: 73pp: English

The present invention describes aggregates of certain proteins, binding molecule known as an inti-

XX
CC
CC
The present invention provides the protein and coding sequences of two human nuclear factors associated with dsRNA (NFAR-1 and NFAR-2). These are transcriptional regulators and are thought to play a role in apoptosis and tumourigenesis. The coding sequence (found on chromosome 19p13.1-13.2) is useful as a probe to detect rearrangements in tumour cells and the protein is useful for determining structure-function relationships.

xx
SO Sequence 145 AA;

Query	Match	13.9%	Score 187.5;	DB 22;	Length 894;
Best Local Similarity		26.0%	Pred. No.	1.3e-07;	
Matches	72;	Conservative	33;	Mismatches	125;
				Indels	47;
				Gaps	13;

Fri Feb 1 20:20:31 2002

us-09-591-632-50_copy_153_405.rag

				GenCore version 4.5
Copyright (c) 1993 - 2000				Compugen Ltd.
OM protein - protein search, using sw model				
Run on:	February 1, 2002, 15:01:55			; Search time 37.66 Seconds (without alignments)
Scoring table:	BLOSUM62			477.735 Million cell updates/sec
Gapop 10.0 , Gapext 0.5				
Searched:	473505 seqs, 1462/2329 residues			
Total number of hits satisfying chosen parameters:	473505			
Minimum DB seq length:	0			
Maximum DB seq length:	2400000000			
Post-processing:	Minimum Match 0%			
	Maximum Match 100%			
Database :	SP TREMBL_17:*			
	1: sp_archaea:*			
	2: sp_bacteria:*			
	3: sp_fungi:*			
	4: sp_human:*			
	5: sp_invertebrate:*			
	6: sp_mammal:*			
	7: sp_mic:*			
	8: sp_organelle:*			
	9: sp_phage:*			
	10: sp_plant:*			
	11: sp_frodet:*			
	12: sp_virus:*			
	13: sp_vertebrate:*			
	14: sp_unclassified:*			
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.				
SUMMARIES				
Result No.	Score	Query Match Length	DB ID	Description
1	701	100.0	685	3 O9HGVI
2	695	99.1	224	3 O07089
3	229.5	32.7	712	3 Q9HGIV
4	228	32.5	299	3 Q0UVK3
5	212	30.2	428	3 Q0URS3
6	212	30.2	700	3 Q9HGIB
7	204	29.1	320	3 Q0UVK2
8	194.5	27.7	701	3 Q9HGIV
9	190	27.1	662	3 Q9HGIV
10	177	25.2	274	3 Q0UVJ9
11	175	25.0	742	3 Q9HGIV
12	170	24.3	200	1 Q9D036
13	167.5	23.9	305	3 Q0UVKO
14	162	23.1	2382	5 Q9B119
15	160	22.8	407	3 Q9A079
16	152.5	21.8	215	3 Q03117
17	149	21.3	456	5 Q0936
18	148	21.1	419	10 Q94942
19	147.5			409
20				5 Q19414
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				
84				
85				
86				
87				
88				
89				
90				
91				
92				
93				
94				
95				
96				
97				
98				
99				
100				
101				
102				
103				
104				
105				
106				
107				
108				
109				
110				
111				
112				
113				
114				
115				
116				
117				
118				
119				
120				
121				
122				
123				
124				
125				
126				
127				
128				
129				
130				
131				
132				
133				
134				
135				
136				
137				
138				
139				
140				
141				
142				
143				
144				
145				
146				
147				
148				
149				
150				
151				
152				
153				
154				
155				
156				
157				
158				
159				
160				
161				
162				
163				
164				
165				
166				
167				
168				
169				
170				
171				
172				
173				
174				
175				
176				
177				
178				
179				
180				
181				
182				
183				
184				
185				
186				
187				
188				
189				
190				
191				
192				
193				
194				
195				
196				
197				
198				
199				
200				
201				
202				
203				
204				
205				
206				
207				
208				
209				
210				
211				
212				
213				
214				
215				
216				
217				
218				
219				
220				
221				
222				
223				
224				
225				
226				
227				
228				
229				
230				
231				
232				
233				
234				
235				
236				
237				
238				
239				
240				
241				
242				
243				
244				
245				
246				
247				
248				
249				
250				
251				
252				
253				
254				
255				
256				
257				
258				
259				
260				
261				
262				
263				
264				
265				
266				
267				
268				
269				
270				
271				
272				
273				
274				
275				
276				
277				
278				
279				
280				
281				
282				
283				
284				
285				
286				
287				
288				
289				
290				
291				
292				
293				
294				
295				
296				
297				
298				
299				
300				
301				
302				
303				
304				
305				
306				
307				
308				
309				
310				
311				
312				
313				
314				
315				
316				
317				
318				
319				
320				
321				
322				
323				
324				
325				
326				
327				
328				
329				
330				
331				
332				
333				
334				
335				
336				
337				
338				
339				
340				
341				
342				
343				
344				

KW	GTP-binding; Protein biosynthesis.	GN	SUP35.
SQ	SEQUENCE 701 AA; 77227 MW; 37517B34289F554 CRC64;	OS	Zygosaccharomyces rouxii (Candida mogii).
Query Match	27.7%; Score 194.5; DB 3; Length 701;	OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Best Local Similarity	38.5%; Pred. No. 2.8e-11;	OC	Saccharomycetales; Saccharomycetaceae; Zygosaccharomyces.
Matches	60; Conservative 8; Mismatches 31; Indels 57; Gaps 12;	OX	NCBI_TAXID=4956;
QY	1 MSDSNQGN-----NQONYQQYSQNGNQQQCNRNYQGYQAYNAQAOA 42	RN	[1]
Db	1 MSDDQQRNQDKLSQDFONTSIGSGEQQQSYQQYQQ--QPOQNN-----FNAANSAPT 50	RP	SEQUENCE FROM N.A.
QY	43 -----GGYYQNTQ-GSSGYQGGYQDNPDAGYQDQYQDNPDGQYQY-NPGGGY 88	RT	Santoso A., Chien P., Weissman J.S.;
Db	51 FTBPSGQOGQGQSGCYQGGYQGGYQONYSQGYQNYN-OQYQNYQ-QYQYQNNRGY 105	RT	"Molecular Basis of a Yeast Prion Species Barrier.";
QY	89 QQQFNPOGGRGIYKFNFYNNNNLQGYQAGFOP-QSQG 123	RL	Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
Db	106 -NSYNNRGGYNNNNYNN-----QdQDQPVQNQ 132	CC	-I- SIMILARITY: TO GTP-BINDING ELONGATION FACTOR FAMILY.
RESULT	9	DR	EMBL; AF206292; AAC140071; -.
Q9HGI4		DR	InterPro: IPRO00705; GTP_EFTU.
ID	Q9HGI4 PRELIMINARY; PRT; 662 AA.	DR	PF00009; GTP_EFTU; 1.
AC	O9HGI4; 01-MAR-2001 (TREMBrel. 16, Created)	KW	GTP-binding; protein biosynthesis.
DT	01-JUN-2001 (TREMBrel. 17, Last sequence update)	FT	NON-TER
DT	01-JUN-2001 (TREMBrel. 17, Last annotation update)	SEQUENCE	274 AA; 30173 MW; 9ECE2F75C8EC4F11 CRC64;
DE	POLYPEPTIDE RELEASE FACTOR 3.	RESULT	25.2%; Score 177; DB 3; Length 274;
GN	SUP35.	Best Local Similarity	44.8%; Pred. No. 5.4e-10;
OS	zygosaccharomyces rouxii (Candida mogii).	Matches	56; Conservative 3; Mismatches 32; Indels 34; Gaps 10;
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;	RN	QY	1 MSDSNQGN-----NQONYQQYSQNGNQQQCNRNYQGYQAYNAQAOA 42
Saccharomycetales; Saccharomycetaceae; Zygosaccharomyces.	RN	Db	1 MSDPNQ-NQQ-----QGQQNQAGGNYQQYFQLKLNQAOQGGG-YQFYGGYAGYQY 51
[1]	RN	QY	59 GYQOYNPDGYQDQYQDNPDAGYQDQYQDNPDGQY-QOFNPQGGGRGYKNEFYNNNLQGYQAF 117
PRAM; PRO0009; GTP_EFTU; 1.	RN	Db	52 GYQ-----GYQQY--QGQQP--QQCGYQYQDAGQGLNNLGPS----- 89
PRAM; PRO0015; ELONGATINFCT.	RN	QY	118 QPQSQ 122
GTP-binding; Protein biosynthesis.	RN	Db	90 QPQQQ 94
SEQUENCE 662 AA; 73780 MW; 2641046199FB1E04 CRC64;	RN	RESULT	11
RN	SEQUENCE FROM N.A.	Q9HGI5	25.2%; Score 177; DB 3; Length 274;
RA	Nakayashiki T., Ebihara K., Nakamura Y.;"	ID	Best Local Similarity
RT	"PSI-like properties of fungus Sup35 Proteins."	Q9HGI5	44.8%; Pred. No. 5.4e-10;
RL	Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.	AC	Matches
CC	-I- SIMILARITY: TO GTP-BINDING ELONGATION FACTOR FAMILY.	DE	32; Indels 34; Gaps 10;
DR	AB039753; BAB12684.2; -.	GN	QY
DR	InterPro: IPR000795; GTP_EFTU.	GT	1 MSDSNQGN-----NQONYQQYSQNGNQQQCNRNYQGYQAYNAQAOA 42
DR	PRAM; PRO0009; GTP_EFTU; 1.	DR	1 MSDPNQ-NQQ-----QGQQNQAGGNYQQYFQLKLNQAOQGGG-YQFYGGYAGYQY 51
PRAM; PRO0015; ELONGATINFCT.	RN	QY	59 GYQOYNPDGYQDQYQDNPDAGYQDQYQDNPDGQY-QOFNPQGGGRGYKNEFYNNNLQGYQAF 117
GTP-binding; Protein biosynthesis.	RN	Db	52 GYQ-----GYQQY--QGQQP--QQCGYQYQDAGQGLNNLGPS----- 89
SEQUENCE 662 AA; 73780 MW; 2641046199FB1E04 CRC64;	RN	RESULT	11
RN	SEQUENCE FROM N.A.	Q9HGI5	25.2%; Score 177; DB 3; Length 274;
RA	Nakayashiki T., Ebihara K., Nakamura Y.;"	ID	Best Local Similarity
RT	"PSI-like properties of fungus Sup35 Proteins."	Q9HGI5	44.8%; Pred. No. 5.4e-10;
RL	Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.	AC	Matches
CC	-I- SIMILARITY: TO GTP-BINDING ELONGATION FACTOR FAMILY.	DE	32; Indels 34; Gaps 10;
DR	AB039752; BAB12683.1; -.	GN	QY
DR	InterPro: IPR000795; GTP_EFTU.	GT	1 MSDSNQGN-----NQONYQQYSQNGNQQQCNRNYQGYQAYNAQAOA 42
DR	PRAM; PRO0009; GTP_EFTU; 1.	DR	1 MSDPNQ-NQQ-----QGQQNQAGGNYQQYFQLKLNQAOQGGG-YQFYGGYAGYQY 51
PRAM; PRO0015; ELONGATINFCT.	RN	QY	59 GYQOYNPDGYQDQYQDNPDAGYQDQYQDNPDGQY-QOFNPQGGGRGYKNEFYNNNLQGYQAF 117
GTP-binding; Protein biosynthesis.	RN	Db	52 GYQ-----GYQQY--QGQQP--QQCGYQYQDAGQGLNNLGPS----- 89
SEQUENCE 742 AA; 80125 MW; 3C8B3F5738206C0A CRC64;	RN	RESULT	10
RN	SEQUENCE FROM N.A.	Q9UVJ9	25.0%; Score 175; DB 3; Length 742;
RA	Nakayashiki T., Ebihara K., Nakamura Y.;"	ID	Best Local Similarity
RT	"PSI-like properties of fungus Sup35 Proteins."	Q9UVJ9	35.1%; Pred. No. 2.4e-09;
RL	Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.	AC	Matches
CC	-I- SIMILARITY: TO GTP-BINDING ELONGATION FACTOR FAMILY.	DE	44; Indels 58; Gaps 11;
DR	AB039752; BAB12683.1; -.	GN	QY
DR	InterPro: IPR000795; GTP_EFTU.	GT	1 MSDS-NQG-NNDQNYQQYSQ-NENQQGNNRNYQGYQAYNAQAO----- 40
DR	PRAM; PRO0009; GTP_EFTU; 1.	DR	1 MSDQFNQDFANKANINEQGOGOGQGQG---OQYQQAPOQOFVPGQSFPVGQSFPVGQ 56
PRAM; PRO0015; ELONGATINFCT.	RN	QY	41 --PAGGYQNTQGYSGYQ-----OGGYQQYNP-----DAGYQQYIN 75
GTP-binding; Protein biosynthesis.	RN	Db	97 PQQQ 100

Db	57 SFAPGGYDG-YQSGGGYDGGYQGGYQGNNRGYSGRGRGGYRGGYOGGYNQGN 115	DE SUP35 HOMOLOG (FRAGMENT).
Qy	76 QGGYQYQNPQGGYQQFNPOGGRGNYKNFVNINNLOGYQA 115	GN SUP35.
ID	Q9DUVK0	OS Saccharomyces ludwigii.
AC	Q9DUVK0;	RT "Molecular Basis of a Yeast Prion Species Barrier."
DT	01-MAY-2000 (TREMBLrel. 13, Created)	RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)	CC {- SIMILARITY: TO GTP-BINDING ELONGATION FACTOR FAMILY.
DE	2610111C21Rik PROTEIN.	DR EMBL: AF206291; AAFL1006.1; -.
GN		DR InterPro: IPR00795; GTP_BEFU.
OS	Mus musculus (Mouse)	DR PFM0009; GTP_EFU; 1.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus. OX NCBI_TaxID:36035;	KW GTP-binding; Protein biosynthesis.
OX	NCBI_TaxID:10090;	FT NON_TER 305 305 MW; 57258C4B3634EES CRC64;
RN	[1]	SQ SEQUENCE FROM N.A.
RP		RA Santoso A., Chien P., Weissman J.S.;
RC		RA "Molecular Basis of a Yeast Prion Species Barrier."
RX	MEDLINE-21085660; PubMed=1121851;	RA Saccharomyces ludwigii.
RX	Kawai J., Shingawa A., Shibata K., Yoshino M., Ichii M., Imai Y., Akawa T., Hara A., Fukunii H., Konno H., Adachi J., Fukuda S., Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamoto I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehnl P., Lewis S., Matsuo Y., Nikaido T., Pasolli G., Quackenbush J., Schreinl L., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Bustindich S., Hill D., Hofmann M., Hume D.A., Kamila M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez T., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;	RA Saccharomyces ludwigii.
RX	"Functional annotation of a full-length mouse cDNA collection.";	RA Saccharomyces ludwigii.
RT	Nature 409:685-690(2001); EMBL; AK011843; BAB27875.1; -	RA Saccharomyces ludwigii.
RL	409:685-690(2001); EMBL; AK011843; BAB27875.1; -	RA Saccharomyces ludwigii.
DR	MGD; MGI:1917689; 2610111C21Rik	RA Saccharomyces ludwigii.
SO	SEQUENCE 200 AA; 21956 MW; A78C141409044215 CRC64;	RA Saccharomyces ludwigii.
Query Match	24.3%; Score 170; DB 11; Length 200;	RESULT 14
Best Local Similarity	30.6%; Pred. No. 1.9e-09;	Query Match 23.9%; Score 167.5; DB 3; Length 305;
Matches	49; Conservative 9; Mismatches 46; Indels 56; Gaps 7;	Best Local Similarity 41.0%; Pred. No. 5.1e-09; Matches 48; Conservative 5; Mismatches 41; Indels 23; Gaps 7;
Qy	1 MSD---SNGNNNQHQOQSNGNQOOGNNRNYQGQAYNAQOPIAGGYYONQGSGYQ 56	Qy 5 NOGNNNDQYNS---QNSQGQGNRNYQGQAYNAQOPIAGGYYONQGSGYQ 58
Db	1 MSDGSYSQSQSGEQSSYSSIGNQSGQGQMP-QSYSGYQGTTD--SSYQYNGYGSQG 57	Db 18 NVAGGRGQYQINAGDQLQDQOQOOQOGYQAOYQNAQSOOG--YQAYQYTAQOQQ 75
Qy	57 OG-----GYQO-----GYQO-----GYQO-----GYQO-----GYQO-----GYQ 75	Qy 59 GQQYIN----PDAGYQ----QYQ----PQ----GGYQYNGQGQYQOFNQGGRN 100
Db	58 QNQSGYSQSQSGYSENOKQSSQGQSYNQGQNTESGGQGGRAPSQGSDYQGQDSYDQ 117	Db 76 GYQAYQYTAQOQQGQYQAHQYQNMGVQPOQOAVYQGQYNAQPOQAVSLDYNSKQ 132
RESULT	13	RESULT 12
O9UVK0	PRELIMINARY; PRT; 305 AA.	RP SEQUENCE FROM N.A.
ID	Q9UVK0	RA Santoso A., Chien P., Weissman J.S.;
AC	Q9UVK0;	RA "Molecular Basis of a Yeast Prion Species Barrier."
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	RA Saccharomyces ludwigii.
DT	01-MAY-2000 (TREMBLrel. 17, Last annotation update)	RA Saccharomyces ludwigii.
Qy	76 QGGYQYQNPQGGYQQFNPOGGRGNYKNFVNINNLOGYQA 115	RA Saccharomyces ludwigii.
Db	118 QSGYDQH--QGSYDEQSINYQ----QHDSYNQNOQSYHS 149	RT "Graal a Drosophila gene coding for several mosaic serine proteases.";
Qy		RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
Qy		DR EMBL: AJ309005; CAC35209.1; -.
Db		DR SIGNAL. 1
Qy		FT POTENTIAL_SIGNAL. 1
Db		SEQUENCE 2382 AA; 264348 MW; 51C85282B0683'D4 CRC64;
Query Match	23.1%; Score 162; DB 5;	Query Match 23.1%; Score 162; DB 5;
Best Local Similarity	31.0%; Pred. No. 1.5e-07;	Length 2382;
Matches	52; Conservative 11; Mismatches 55; Indels 50; Gaps 10;	Matches 52; Conservative 11; Mismatches 55; Indels 50; Gaps 10;
Qy	4 SNQGNNQO---NTO---QYSQNG---NQOQGNRNYQ-----GIA-----YNAQAOPIAG 43	Qy 4 SNQGNNQO---NTO---QYSQNG---NQOQGNRNYQ-----GIA-----YNAQAOPIAG 43
Db	1029 ISLDGGGNOPPGKYQAPYIISNGGIFTPSQNNGGQIAPFQONNGGQPSQGNG 1088	Db 1029 ISLDGGGNOPPGKYQAPYIISNGGIFTPSQNNGGQIAPFQONNGGQPSQGNG 1088
Qy	44 GYQNYQG-YSCYQGQYQYQNPQDGQYQQWNPQGGYQ-----YNP 84	Qy 44 GYQNYQG-YSCYQGQYQYQNPQDGQYQQWNPQGGYQ-----YNP 84
Db	1089 GYTPYQNPFEQYHQPQNOSSYQNTGGYQPPLLPHGGYQPPSOTYGGFOQPNRTYEDFHNP 1148	Db 1089 GYTPYQNPFEQYHQPQNOSSYQNTGGYQPPLLPHGGYQPPSOTYGGFOQPNRTYEDFHNP 1148

Db 1149 SQTVGFFQPPYSWKGGYQPHNQLNGGYNQPNOSSSFQA - NGGYQPPSQ 1195

RESULT	15
ID	094079
AC	094079;
PRELIMINARY;	PRT;
PT	407 AA.
DT	01-MAY-1999 (TREMBLrel. 10, Created)
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	YEAST (S.CEREVISEIAE) FUS1 (FUSION GENE) AND BIK1 GENE REGION.
OS	RNQ1 AND YCL028W
CC	Saccharomyces cerevisiae (Baker's Yeast).
CC	Eukarya; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX	NCBL-TAXID=4932;
RN	[1]
RP	SEQUENCE FROM N.A.
FX	MEDLINE=8728684; PubMed=3302672;
RA	Trueheart J., Boeke J.D., Fink G.R.;
RT	"Two genes required for cell fusion during yeast conjugation: evidence for a pheromone-induced surface protein.";
RT	Mol. Cell. Biol. 7:2316-2328(1987).
RL	EMBL: M16717; AAA36151.1; -.
DR	SGD: S000533; RNQ1.
SQ	SEQUENCE 407 AA; 43262 MW; 6B26940E74A804C3 CRC64:
Query Match	22.8%; Score 160; DB 3; Length 407;
Best Local Similarity	38.1%; Pred. NO. 3.7e-08;
Matches	51; Conservative 5; Mismatches 34; Indels 44; Gaps 54
Qy	SDSNQNGNNQNYOQYSONGNQQGNNRNYQGYQAYNAQAOQPAAGYYQNYQGYSG-----
Db	219 NNQNNSNNSQGQYNGSYQNGNQNS---OQYN---NQTYQOGNGGYQQQOGQSGGAFSSLA 271
Qy	55 -----YQQGQYQNPDAQYQQQNPQGGYQQNPQGGYQQNPQGRGRGNKFNENNN 109
Db	272 SMAQSGLGSGQTOQN----QQDYNQOQ---QNN----QDQYQOQGQ-----NYQHQ 311
Qy	110 LQGIOQAGFQBSQG 123
Db	312 QQG-----QQQQG 320

Search completed: February 1, 2002, 15:07:10
Job time: 315 sec

			GenCore version 4.5
Copyright (c) 1993 - 2000			Compugen Ltd.
Run on:	February 1, 2002, 15:02:20	; Search time 15.38 Seconds	(without alignments)
			293.223 Million cell updates/sec
Title:	US-09-591-632-2_COPY_1_123		
Perfect score:	701		
Sequence:	1 MSDSNGNNQNYQQYSONG.....	FNYYNNNLQGYQAGFOPQSOG	123
Scoring table:	BLOSUM62		
	Gapop 10.0 , Gapext 0.5		
Searched:	100059 seqs, 36664827 residues		
Total number of hits satisfying chosen parameters:	100059		
Minimum DB seq length:	0		
Maximum DB seq length:	200000000		
Post-processing:	Minimum Match 0%		
	Maximum Match 100%		
	Listing first 45 summaries		
Database :	SwissProt_39;*		
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.			
		SUMMARIES	
Result No.	Score	Query Match Length DB ID	Description
1	701	100.0	ERF2_YEAST
2	228	32.5	ERF2_CANAL
3	167	23.8	R856_HUMAN
4	166	23.7	YCC8_YEAST
5	156	22.4	ERF2_PICPI
6	157	21.0	FUS_BOVIN
7	147	20.4	ANX7_DICDI
8	143	20.4	PTEN_MCGE
9	143	20.4	FUS_MOUSE
10	141	20.1	SSXT_HUMAN
11	139	19.8	SSXT_MOUSE
12	136	19.4	FUS_HUMAN
13	135	19.3	YBM6_YEAST
14	132	18.8	TF2_BORBU
15	128	18.3	ROD_RAT
16	127	18.1	ROD_HUMAN
17	126.5	18.0	YRF4_YEAST
18	123	17.5	YR53_CABEV
19	122.5	17.5	CBBB_BACTV
20	121	17.3	YXNA_RUMFL
21	118	16.8	GTT4_WHEAT
22	116	16.5	34KD_MCPA
23	115.5	16.5	8261_CRA_BACUF
24	115	16.4	GTT5_WHEAT
25	115	16.4	DPL1_SCHPO
26	114.5	16.3	EG11_CANAL
27	114	16.3	LGG3_CANFA
28	113.5	16.2	SYPH_BOVIN
29	113	16.1	CRA_BACUF
30	112	16.0	YSA1_CABEL
31	112	16.0	OVO_DROME
32	111.5	15.9	P1512_drosophila
33	111	15.8	PO8247_homo_sapien
34			YSO5_CAEEL
35			R021_XENLA
36			SMPL_HUMAN
37			CUP9_DROME
38			AGO1_ARATH
39			YK03_CAEEL
40			YKT3_CAEEL
41			GLT3_WHAT
42			SOD_DROME
43			GLT0_WHEAT
44			GRPA_MEDEFA
45			ENS_MOUSE

ALIGNMENTS			
RESULT	1		
ID	ERF2_YEAST	STANDARD;	PRT; 685 AA.
AC	P05453; P05420;		
DT	01-NOV-1988 (Rel. 09, Created)		
DT	01-NOV-1988 (Rel. 09, Last sequence update)		
DT	15-JUL-1998 (Rel. 36, Last annotation update)		
DE	EUKARYOTIC PEPTIDE CHAIN RELEASE FACTOR GTP-BINDING SUBUNIT (ERF2), TRANSLATION RELEASE FACTOR 3 (ERF3) (OMNIPOTENT SUPPRESSOR PROTEIN 2) (G1 TO S PHASE TRANSITION PROTEIN 1).		
DE	SUP35 OR SUP2 OR SUF12 OR GST1 OR SAL3 OR PNMD2 OR YDR172W OR YD9395_05.		
OS	Saccharomyces cerevisiae (Baker's yeast); Eukaryota; Fungi; Ascomycota; Saccharomycetales; Saccharomycetaceae; Saccharomyces.		
OC	NCBI TaxID=4932;		
OX	[1]		
RA	Kushnirov V.V., Ter-Avanesyan M.D., Telckov M.V., Surguchov A.P., Smirnov V.N., Inge-Vechtomov S.G.,		
RA	"Nucleotide sequence of the SUP2 (SUP35) gene of Saccharomyces cerevisiae"; Gene 66:45-54(1988).		
RA	SEQUENCE FROM N.A.		
RA	MEDLINE#87219095; PubMed=3556215;		
RA	Kushnirov V.V., Ter-Avanesyan M.D., Surguchov A.P., Smirnov V.N., Inge-Vechtomov S.G.; "Localization of possible functional domains in sup2 gene product of the yeast Saccharomyces cerevisiae.", FEBS Lett. 215:257-260(1987).		
RA	[3]		
RA	SEQUENCE FROM N.A.		
RA	MEDLINE#88172503; PubMed=3280807;		
RA	Wilson P.G., Culbertson M.R.; "SUF12 suppressor protein of yeast: A fusion protein related to the EF-1 family of elongation factors.", J. Mol. Biol. 199:559-573(1988).		
RA	[4]		
RA	SEQUENCE FROM N.A.		
RA	MEDLINE#88296422; PubMed=2841115;		
RA	Kukuchi Y., Shimatake H., Kikuchi A.; "A yeast gene required for the G1-to-S transition encodes a protein containing an A kinase target site and GTPase domain.", EMBO J. 7:1175-1182(1988).		
RA	[5]		
RA	SEQUENCE FROM N.A.		
RA	STRAYN-S28BC / AB972; Murphy L., Harris D.E., Barrell B.G., Rajandream M.A.; Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.		
RA	[6]		
RA	FUNCTION.		
RA	MEDLINE#9616209; PubMed=7556078;		
RA	Stansfield I., Jones K.M., Kushnirov V.V., Dragkesamanskaya A.R., Poznyakovski A.I., Paushkin S.V., Nierras C.R., Cox B.S.,		

RL	Genomics 38:51-57(1996).	DR	SMART; SM00547; ZnF_RBZ; 1.
RN	[2]	DR	PROSITE; PS5102; RRM; 1.
RP	SEQUENCE FROM N.A. (SHORT ISOFORM), AND SEQUENCE OF 282-297 & 307-320.	DR	PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
RX	MEDLINE-97045110; PubMed-8890175;	KW	Proto-oncogene; RNA-binding; DNA-binding; Nuclear protein; Repeat;
RA	Bertolotti A., Lutz J., Heard D.J., Chambon P., Tora L.;	KW	Alternative splicing; Chromosomal translocation; Zinc-finger; Zinc;
RT	"TAFI116B, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TIS/FUS and EWS is associated with both TFIID and RNA polymerase II";	KW	Metal-binding.
RT	RT	FT	DOMAIN 1 208 GIN/GLN/SER/TYR RICH.
RL	EMBO J. 15:5022-5031(1995).	FT	DOMAIN 234 320 RNA-BINDING (RRM).
RN	[3]	FT	DOMAIN 320 590 ARG/GLY-RICH.
RP	SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).	FT	ZNFING 360 379 C4-TYPE (POTENTIAL).
RX	MEDLINE-99013073; PubMed-9795213;	FT	DOMAIN 407 575 21 X APPROXIMATE TANDEM REPEATS OF D-R-[S,G](O,3)-G-G-Y-G-G.
RA	Morohoshi F., Ootsuka Y., Arai K., Ichikawa H., Mitani S.,	FT	REPEAT 407 413 1.
RA	Munakata M., Ohki M.;	FT	REPEAT 414 420 2.
RT	"Genomic structure of the human RBP56/ttaFI16B and FUS/TIS genes.";	FT	REPEAT 421 429 3.
RL	Gene 221:191-198(1998).	FT	REPEAT 430 439 4.
CC	-I- FUNCTION: RNA AND SSDNA-BINDING PROTEIN THAT MAY PLAY SPECIFIC ROLES DURING TRANSCRIPTION INITIATION AT DISTINCT PROMOTERS. CAN ENTER THE PREINITIATION COMPLEX TOGETHER WITH THE RNA POLYMERASE II (POL II) TRANSCRIPTIONAL MULTIROTEIN COMPLEX, TOGETHER WITH THE TATA-BINDING PROTEIN (TBP) AND OTHER TBP-ASSOCIATED FACTORS (TAF(I)S).	FT	REPEAT 440 448 5.
CC	-I- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).	FT	REPEAT 449 457 6.
CC	-I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG ISOFORM (SHOWN HERE) AND A SHORT ISOPRFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.	FT	REPEAT 458 465 7.
CC	-I- TISSUE SPECIFICITY: UBIQUITOUS. OBSERVED IN ALL FETAL AND ADULT TISSUES.	FT	REPEAT 466 473 8.
CC	-I- DISEASE: A FORM OF EXTRASKELETAL MYOID CHONDROSARCOMAS (EMC) IS CHARACTERIZED BY A CHROMOSOMAL TRANSLOCATION T(9;17)(Q22;Q11) THAT INVOLVES TAF2N AND NKX4A3.	FT	REPEAT 474 481 9.
CC	-I- SIMILARITY: CONTAINS 1 RAMBP-TYPE ZINC FINGER.	FT	REPEAT 482 488 10.
CC	-I- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.	FT	REPEAT 489 496 11.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_licensee@ib-sib.ch).	FT	REPEAT 497 503 12.
CC	FT	REPEAT 504 510 13.
CC	FT	REPEAT 511 517 14.
CC	FT	REPEAT 518 524 15.
CC	FT	REPEAT 525 533 16.
CC	FT	REPEAT 534 543 17.
CC	FT	REPEAT 544 551 18.
CC	FT	REPEAT 552 560 19.
CC	FT	REPEAT 561 568 20.
CC	FT	REPEAT 569 575 21.
CC	FT	REPEAT 592 AA; MISSING (IN SHORT ISOFORM).
CC	FT	VARSPLIC 60 62 MISSING (IN SHORT ISOFORM).
CC	FT	SEQUENCE 592 AA; 61830 MW; 73D3YC171E2BCA CRC64;
EMBL	U51134; AAC50332.1; -.	Query	Match 23.8%; Score 167; DB 1; Length 592;
EMBL	X9893; CAA67398.1; -.	Best Local Similarity	29.1%; Pred. No. 8. Be-07;
EMBL	AB00067; BAA33811.1; -.	Matches	46; Conservative 17; Misnatches 45; Indels 50; Gaps 7;
EMBL	AB00057; BAA33811.1; JOINED.	Qy	1 MSDS---NQGNNQNQWNQQYSONGNDQG-----NNRY-OGYAYNAQAPQA 42
EMBL	AB00058; BAA33811.1; JOINED.	Db	1 MSDSGSGYQGSQGSEQSYGQSTYGNPGSQGYQASOSYSGYQITDSSYQONYSGYQSQ 60
EMBL	AB00059; BAA33811.1; JOINED.	Qy	43 GGYYONQYQSYQOGGQQ-YN-----PDAGYQQQINPOG 77
EMBL	AB00060; BAA33811.1; JOINED.	Db	61 SQQSQQGGYENQKQSSYQSYQPNNGQQQINMESSGQGGGRAPSYPDPDYQDPSDQG 120
EMBL	AB00061; BAA33811.1; JOINED.	Qy	78 GQOQYNPOGGQQQFNPOGGGRNYKNFYNNNLQGYQA 115
EMBL	AB00062; BAA33811.1; JOINED.	Db	121 GYDQH-QGSVDEQSN----YDOQHDSYSQNOQSYHS 151
EMBL	AB00063; BAA33811.1; JOINED.	RESULT	4
EMBL	AB00064; BAA33811.1; JOINED.	YCC8 YEAST	STANDARD; PRY; 405 AA.
EMBL	AB00065; BAA33811.1; JOINED.	ID	YCC8 YEAST
EMBL	AB00066; BAA33811.1; JOINED.	AC	P25367;
EMBL	AB00067; BAA33812.1; -.	DT	01-MAY-1992 (Rel. 22, Created)
EMBL	AB00057; BAA33812.1; JOINED.	DT	01-MAY-1992 (Rel. 22, Last sequence update)
EMBL	AB00058; BAA33812.1; JOINED.	DT	15-DEC-1998 (Rel. 37, Last annotation update)
EMBL	AB00059; BAA33812.1; JOINED.	DT	DE HYPOTHETICAL 42.6 KDA PROTEIN IN BIK1-FUS1 INTERGENIC REGION.
EMBL	AB00061; BAA33812.1; JOINED.	GN	YCL028W OR YCL28W OR YCL181.
EMBL	AB00062; BAA33812.1; JOINED.	OS	Saccharomyces cerevisiae (Baker's yeast).
EMBL	AB00053; BAA33812.1; JOINED.	OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
EMBL	AB00054; BAA33812.1; JOINED.	OX	Saccharomycetales; Saccharomycetaceae; Saccharomyces;
EMBL	AB00065; BAA33812.1; JOINED.	RN	NCBI_TaxID=493;
MM	601574; -.	Qy	[1]
MM	Interpro; IPR000504; RRM.	RP	SEQUENCE FROM N.A.
IPR000504; IPR000504; Znf_RanBP.	MM	RX	MEDLINE-9137317; PubMed-1897318;
PFam; PF00076; rrm; 1.	MM	RA	Rad M.R., Luetzenkirchen K., Xu G., Kleinhans U., Hollenberg C.P.;
SMART; SM00360; RRM; 1.	MM	RT	"The complete sequence of a 11.953 bp fragment from CIG on chromosome III encompasses four new open reading frames."

RL Yeast 7:533-538(1991).
 RN [2]
 SEQUENCE FROM N.A.
 RP MEDLINE=87286864; PubMed=3302672;
 RA Trueheart J.J., Boeke J.D.; Fink G.R.;
 RT "Two genes required for cell fusion during yeast conjugation: evidence for a pheromone-induced surface protein.";
 RL Mol. Cell. Biol. 7:2316-2320(1987).
 CC - CAUTION: REF 2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 323
 CC ONWARD DUE TO A FRAMESHIFT.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M16717; -, NOT_ANNOTATED_CDS.
 DR X59720; CAAK2357.1; -.
 DR PIR; C2061; C2061.
 DR PIR; S19355.
 DR PIR; S17471.
 DR SGD; S0000533; RNO1.
 KW Hypothetical protein.
 FT DOMAIN 150 405 ASN/GLN-RICH.
 FT CONFLICT 181 181 A -> T (IN REF. 2).
 SQ SEQUENCE 405 AA; 42589 MW; 0629C40A7DE9C043 CRC64;

Query Match 23.7%; Score 166; DB 1; Length 405;
 Best Local Similarity 34.4%; Pred. No. 7.4e-07;
 Matches 55; Conservative 7; Mismatches 44; Indels 54; Gaps 9;

Qy 2 SDSSQGNQNQQNQYQYSQNSQGNQQGNNRQGYQAYNAQAOQPGAGYYQNYQGYGSG----- 54
 Db 219 NNQNSNSNSQGYNQSYQNGNQNS-----QGYN-NQQYQOGNGGQYQQQGQSGCAFSSLA 271

Qy 55 -----YQDGGYQOYNPDAGYQOYNPOG--GYQOYNPOG-----GYQOYNPOG---- 96
 Db 272 SMAQSYLGGGQTQSN-----QQYQNCQNCQYQOCQGONQYQHQQCQQQO---QHSS 323

Qy 97 -----GRGVWKFVNNNNLQYQAGFQPOSQ 123
 Db 324 SFSSALASMASSYLGNNNSNSYGGQQQANEGY-RPOHNG 362

RESULT 5
 ERB2_PICPI ID ERB2_PICPI STANDARD; PRT; 741 AA.
 AC P23637; DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE EUKARYOTIC PEPTIDE CHAIN RELEASE FACTOR GTP-BINDING SUBUNIT (ERF2) (TRANSLATION RELEASE FACTOR 3) (ERF3) (ERF-3) (OMNIPOTENT SUPPRESSOR PROTEIN 2).
 DE SUP2.
 OS Pichia pinus (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 RN [1]
 SEQUENCE FROM N.A.
 RC STRAIN=M4;
 RX MIDDLE=M-91181341; PubMed=200663;
 RA Kushnirov V.V., Ter-Avanesyan M.D., Didichenko S.A., Smirnov V.N., Chernoff Y.O., Berkach I.L., Novikova O.N., Inge-Vechtomov S.G., Neistat M.A., Tolstorukov I.I.;
 RT "Divergence and conservation of SUP2 (SUR35) gene of yeast Pichia pinsus and Saccharomyces cerevisiae.", Yeast 6:461-472(1990).

CC -1- FUNCTION: INVOLVED IN TRANSLATION TERMINATION. STIMULATES THE ACTIVITY OF ERF1. BINDS GUANINE NUCLEOTIDES.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC ERF3 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X56910; CAAK0231.1; -.
 DR PIR; S12921; S12921.
 DR HSSP; Q01698; ITTT.
 DR InterPro; IPR00795; GTP_EFTU.
 DR Pfam; PF00009; GTP_EFTU_1.
 DR PROSITE; PS00301; EFACTOR_GTP; 1.
 KW Protein biosynthesis; Repeat; Phosphorylation; GTP-binding.
 FT DOMAIN 162 315 SEVERAL SORT OF REPEATS.
 FT DOMAIN 312 741 CHARGED.
 FT SIMILAR 325 332 TO ELONGATION FACTORS EF-1.
 FT NP_BIND 402 406 GTP (BY SIMILARITY).
 FT NP_BIND 464 470 GTP (BY SIMILARITY).
 FT MOD_RES 399 399 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 741 AA; 82373 MW; A3F30DA3B367EB45 CRC64;

Query Match 23.7%; Score 166; DB 1; Length 741;
 Best Local Similarity 37.5%; Pred. No. 1.3e-06;
 Matches 45; Conservative 7; Mismatches 36; Indels 32; Gaps 7;

Qy 11 QNYQYQYSQNGNQQGNNRQGYQAYNAQAOQPGAGYYQNYQGYGQYQNPDAQYQ 70
 Db 56 QQEQQFGQGQQQONYQ-GGYNNYNR--GGYNNRGYNNNSRGYNSN--SYN 107

Qy 71 QPYRQPGGYQYQNFQGGQQQFNPGGRNRYKNFNYNN--LOGYQAGF----QPOSQ 122
 Db 108 TNSN-OGGSYNN-----NYYANNSYNNNNNNNNNNYQGYNNNSQPOQQ 151

RESULT 6
 FUS_BOVIN ID FUS_BOVIN STANDARD; PRT; 512 AA.
 AC Q2809;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE RNA-BINDING PROTEIN FUS (PIGPEN PROTEIN).
 GN FUS.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 RN [1]
 RN NCBI_TaxID=9913;
 RP SEQUENCE FROM N.A.
 RC TISSUE=aorta;
 RX MEDLINE=96175600; PubMed=8631501;
 RA Allegro M.C.; Allegro M.A.;
 RT "A nuclear protein regulated during the transition from active to quiescent phenotype in cultured endothelial cells.", Dev. Biol. 174:288-297(1996).
 RL
 [2]
 RP CARBOHYDRATE BINDING DOMAIN.
 RX MEDLINE=20160719; PubMed=10694442;
 RA Allegro M.C.; Allegro M.A.;
 RT "A C-terminal carbohydrate-binding domain in the endothelial cell regulatory protein, Prgen: new function for an EWS family member.", Exp. Cell Res. 255:270-277(2000).

							RP	SEQUENCE FROM N.A.
CC	-!- FUNCTION: BINDS BOTH SINGLE-STRANDED AND DOUBLE-STRANDED DNA AND PROMOTES ATP-DEPENDENT ANNEALING OF COMPLEMENTARY SINGLE-STRANDED DNA.	RC	STRAINAX2;				RX	MEDLINE=91198150; PubMed=1826615;
CC	SIMILARITY: STRANDED DNAs AND D-LOOP FORMATION IN SUPERHELICAL DOUBLE-STRANDED DNA. MAY PLAY A ROLE IN MAINTENANCE OF GENOMIC INTEGRITY (BY SIMILARITY).	RA	Greenwood M., Tsang A.;				RA	"Sequence and expression of annexin VII of Dictyostelium discoideum.";
CC	-!- SUBUNIT: COMPONENT OF NUCLEAR RIBOPROTEIN COMPLEXES (BY SIMILARITY).	RL	Biochim. Biophys. Acta 1088:429-432(1991).				RN	[2]
CC	-!- SUBCELLULAR LOCATION: NUCLEAR, EXHIBITS DIFFUSE STAINING THROUGHOUT (EXCLUDING NUCLEOLI), TOGETHER WITH A SMALL NUMBER OF INTENSELY STAINED FOCAL POINTS, OR GRANULES, AND PUNCTATE STAINING ALONG THE NUCLEAR ENVELOPE.	RX	Medline=91373375; PubMed=1832674;				RX	[3]
CC	-!- DOMAIN: THE C-TERMINAL DOMAIN BINDS CARBOHYDRATES.	RA	Doring V., Schleicher M., Noegel A.A.;				RA	"Dictyostelium annexin VII (synexin). cDNA sequence and isolation of a gene disruption mutant.";
CC	-!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).	RT	"Dictyostelium annexin VII (synexin). cDNA sequence and isolation of a gene disruption mutant.";				RL	J. Biol. Chem. 266:17509-17515(1991).
CC	-!- SIMILARITY: BELONGS TO THE TEI FAMILY OF RNP PROTEINS.	RN	"Identification of a homologue for annexin VII (synexin) in dictyostelium discoideum";				RN	[4]
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	RX	"Identification of a homologue for annexin VII (synexin) in dictyostelium discoideum";				RX	X-RAY CRYSTALLOGRAPHY (2.45 ANGSTROMS) OF 108.462.
CC	EMBL: U26024; AAC1343.1; -.	RA	Noegel A.A., Jacob U.;				RA	Medline=91107669; PubMed=1824843;
DR	HSSP: P09511; IHA1.	RT	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";				RT	Medline=97375585; PubMed=97331902;
DR	InterPro: IPR000504; RRM.	RT	J. Mol. Biol. 270:79-88(1997).				RL	Medline=97375585; PubMed=97331902;
DR	InterPro: IPR001876; Znf-RanBP.	CC	-!- FUNCTION: CALCIUM-PHOSPHOLIPID-BINDING PROTEIN WHICH PROMOTES MEMBRANE FUSION AND IS INVOLVED IN EXOCYTOSIS.				CC	Medline=97375585; PubMed=97331902;
DR	Pfam: PF00076; rrm; 1.	CC	-!- SIMILARITY: BELONGS TO THE ANNEXIN FAMILY.				CC	Medline=97375585; PubMed=97331902;
DR	Pfam: PF00641; zf-RanBP; 1.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				CC	Medline=97375585; PubMed=97331902;
DR	SMART: SMM0360; RRM; 1.	CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
DR	PROSITE: PS50102; RRM; 1.	CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
DR	RNA-binding; DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;	CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
FT	Metal-binding.	CC	-!- FUNCTION: CALCIUM-PHOSPHOLIPID-BINDING PROTEIN WHICH PROMOTES MEMBRANE FUSION AND IS INVOLVED IN EXOCYTOSIS.				CC	-!- FUNCTION: CALCIUM-PHOSPHOLIPID-BINDING PROTEIN WHICH PROMOTES MEMBRANE FUSION AND IS INVOLVED IN EXOCYTOSIS.
DOMAIN	1 164 GLN/GLY/SER/TYR-RICH.	CC	-!- SIMILARITY: BELONGS TO THE ANNEXIN FAMILY.				CC	-!- SIMILARITY: BELONGS TO THE ANNEXIN FAMILY.
FT	DOMAIN 165 253 GLY-RICH.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
FT	DOMAIN 271 357 RMA-BINDING (RRM).	CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
FT	DOMAIN 357 512 ARG/GLX-RICH.	CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
FT	ZNF-FING 414 433 C4-TYPE (POTENTIAL).	CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
SQ	SEQUENCE 512 AA; 52240 MW; 3653329C04F1386 CRC64;	CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(1997).
		CC	DR	Noegel A.A., Jacob U.;			CC	Noegel A.A., Jacob U.;
		CC	DR	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";			CC	"Crystal structure of the C-terminal tetrad repeat from synexin (annexin VII) of dictyostelium discoideum.";
		CC	DR	J. Mol. Biol. 270:79-88(1997).			CC	J. Mol. Biol. 270:79-88(

RESULT	8	QY	108 NNLOGY-QAGFQPO 120
ID	P65H_MYCGE	STANDARD;	PRT; 372 AA.
ID	FUS_MOUSE	STANDARD;	PRT; 518 AA.
AC	P47459;		
DT	01-OCT-1995 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	PROLINE-RICH P65 PROTEIN HOMOLOG.		
GN	MG217.		
OS	Mycoplasma genitalium.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscetes;		
OC	Mycoplasmataceae; Mycoplasma.		
OX	NCBI_TAXID=2097;		
RN	SEQUENCE FROM N.A.		
RP	[1] STRAIN=ATCC 33530 / G-37;		
RX	MEDLINE=96026346; PUBMED=7569993;		
RA	Fraser, G.; Gaynor, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.; Bult, C.J.; Kerlavage, A.R.; Sutton, G.; Kelley, J.M.; Fritchman, J.L.; Weidman, J.F.; Small, K.V.; Sandusky, M.; Fuhrmann, J.L.; Nguyen, D.T.; Utterback, T.R.; Sauder, D.M.; Phillips, C.A.; Merrick, J.M.; Tomb, J.F.; Dougherty, B.A.; Bott, K.F.; Hu, P.C.; Lucier, T.S.; Peterson, S.N.; Smith, H.O.; Hutchinson, C.A.; III, Venter, J.C.; "the minimal gene complement of Mycoplasma genitalium.", Science 270:397-403(1995).		
RA	-1- SIMILARITY: TO M_PNEUMONIAE PR0MBIN P65.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	-----		
CC	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.		
CC	-1- FUNCTION: BINDS BOTH SINGLE-STRANDED AND DOUBLE-STRANDED DNA AND PROMOTES ATP-INDEPENDENT ANNEALING OF COMPLEMENTARY SINGLE-STRANDED DNAs AND D-LOOP FORMATION IN SUPERHELICAL DOUBLE-STRANDED DNA. MAY PLAY A ROLE IN MAINTENANCE OF GENOMIC INTEGRITY (BY SIMILARITY).		
CC	-1- SUBUNIT: COMPONENT OF NUCLEAR RIBOPROTEIN COMPLEXES (BY SIMILARITY).		
CC	-1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).		
CC	-1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).		
CC	-1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
DR	EMBL; U39701; AAC71436_1; -		
DR	TIGR; MG217; -		
DR	Repeat; Complete proteome.		
FT	DOMAIN 29 128 8 X DP[NQ](Q)FNQ REPEATS.		
FT	REPEAT 29 37 1.		
FT	REPEAT 41 49 2.		
FT	REPEAT 61 69 3.		
FT	REPEAT 81 89 4.		
FT	REPEAT 101 109 5.		
FT	REPEAT 120 128 6.		
FT	REPEAT 139 147 7.		
FT	REPEAT 120 128 8.		
SQ	SEQUENCE 372 AA; 44664 MW; 4C29701D213CE19E CRC64;		
Query Match	20.4%; Score 143; DB 1; Length 372;		
Best Local Similarity	32.7%; Pred. No. 4 8e-05;		
Matches	51; Conservative 12; Mismatches 55; Indels 38; Gaps 8;		
QY	3 DSNQGNQQNQQYQQYQNG---NQQQCN----RYGGYQAYNAQAOQAGGYQNYQCGSG 54		
Db	22 DQNQYQDPNQQQFNQSGFDPNQQFNQPGDFDPNQQYQDQFQDNQQYQD 81		
QY	55 -----YQGGGQYQYNPDAGY-QQYNNPQGGYQYQNNPQGGYQQF--NP 94		
Db	82 PNQQFPNQPGDFPNQQYQDPNQQFNQSGFDPNQQFNQPGDFDPNQQFNQSGFDQDNQQYQD 140		
QY	95 OGGRGNYKNINYNNLNQLOGYAGFO-----PQS 122		
Db	141 NQQDQNPSPDLNQ-QFNQPGFNGSPAFERTPQEQ 175		
RESULT	9	QY	108 NNLOGY-QAGFQPO 120
FUS_MOUSE			
ID	FUS_MOUSE	STANDARD;	PRT; 518 AA.
AC	P5659;		
DT	20-AUG-2001 (Rel. 40, Created)		
DT	20-AUG-2001 (Rel. 40, Last sequence update)		
DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	RNA-BINDING PROTEIN FUS (PIGPEN PROTEIN).		
GN	FUS.		
OS	Mus musculus (Mouse);		
OC	Makarova; Metzoco; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.		
OC	Burdal, C.A.;		
RA	Alapati, S.R.; Zhang, M.; Zhao, X.; Allegro, M.A.; Allegro, M.C.; Burdall, C.A.;		
RA	"Regulation of pigrpen expression in mouse embryos";		
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.		
CC	-1- FUNCTION: BINDS BOTH SINGLE-STRANDED AND DOUBLE-STRANDED DNA AND PROMOTES ATP-INDEPENDENT ANNEALING OF COMPLEMENTARY SINGLE-STRANDED DNAs AND D-LOOP FORMATION IN SUPERHELICAL DOUBLE-STRANDED DNA. MAY PLAY A ROLE IN MAINTENANCE OF GENOMIC INTEGRITY (BY SIMILARITY).		
CC	-1- SUBUNIT: COMPONENT OF NUCLEAR RIBOPROTEIN COMPLEXES (BY SIMILARITY).		
CC	-1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).		
CC	-1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).		
CC	-1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
DR	EMBL; MG1133633; FUS.		
DR	DR; MGI: MGI:002952; Eggshell.		
DR	InterPro; IPR002952; RRM.		
DR	InterPro; IPR001876; Znf-RanBP.		
DR	Pfam; PF00076; rrm; 1.		
DR	DR; Pfam; PF00641; f-RanBP; 1.		
DR	DR; SMART; SM00360; RRM; 1.		
DR	DR; SMART; SM00547; Znf-RBZ; 1.		
DR	DR; PROSITE; PS5002; RRM; 1.		
DR	DR; PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.		
KW	RNA-binding; DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;		
KW	Metal-binding.		
FT	DOMAIN 1 167 GLN/GLY/SER/TYR-RICH.		
FT	DOMAIN 168 265 GLY-RICH.		
FT	DOMAIN 278 364 RNA-BINDING (RRM).		
FT	DOMAIN 364 518 ARG/GLY-RICH.		
FT	ZN_FING 421 440 C4-TYPE (POTENTIAL).		
SQ	SEQUENCE 518 AA; 52673 MW; E06F231BREED/8D6 CRC64;		
Query Match	20.4%; Score 143; DB 1; Length 518;		
Best Local Similarity	29.1%; Pred. No. 6 5e-05;		
Matches	46; Conservative 12; Mismatches 60; Indels 40; Gaps 8;		
QY	2 SDSNQGN---NQNTQYQSONGNGQGNNRQYQYQAYNAQAOQAGGYQNYQCGSQYQG 58		
Db	85 SOSSSQSYGQSSSIYFGYGOOPAPSSTSQSGSSQSSYQPOSGGGYGGQQQSY 144		
QY	59 GYQQ---YNDAGY---QQYNPQ-----GGY---QQYNPQ 85		
Db	145 QGGSQSYNPPQGYQGQONNNSGGGGGGGGNYQODOSMSGGGGGGQDOSGGGG 204		

Db	205 GGYGGQQDRGGGRGGGGYNR---SGGYEPRG 238	Query Match 20.1%; Score 141; DB 1; Length 387;
RESULT	10	Best Local Similarity 35.7%; Pred. No. 7.2e-05;
SSXT_HUMAN	STANDARD; PRT; 387 AA.	Matches 50; Conservative 9; Mismatches 55; Indels 26; Gaps 10;
ID	Q1532; Q16404; 15-DEC-1998 (Rel. 37, created)	Qy 1 MSDSNQNNNQNQNYQQ---SONGNOQ--GNNVYQGQ-AYNQAQPAQG--GYYQNY 49
AC	DT 15-DEC-1998 (Rel. 37, last sequence update)	Db 235 MGQVNQGNHMMGQRQKIPPKRPQOGPPQYSGQEDYQGDQSHGGQGPPECMNQOYPDG 294
DT	20-AUG-2001 (Rel. 40, last annotation update)	Qy 50 QGYSYQGSQWQYQNPADY---QDQYNPQGQYQYNPQGQFNP-OGGRG-NYKNF 104
DE	SSXT PROTEIN (SYNOVIAL SARCOMA, TRANSLOCATED TO X CHROMOSOME) (SYT PROTEIN)	Db 295 NSQYSQGQDQAWQGPBPQGPBPQQOQYPGQGYP-----GQQOQYGPSPGGQYQPNY 348
DE	SSXT OR SYT.	Qy 105 NYNNLNLQGYQ-AGFQPOSQG 123
OS	Homo sapiens (Human)	Db 349 PQG--QGQGYYGYPQPG 365
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCBL_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A., AND SSXT-SX2 FUSION PROTEIN.	
RC	TISSUE-Synovial sarcoma;	
RX	MEDLINE-95038336; PubMed-7951320;	
RA	Clark J., Roques P.J., Crew A.J., Gill S., Shipley J., Chan A.M.-L., Gusterson B.A., Cooper C.S.; Identification of novel genes, SYT and SSX, involved in the t(X;18)(p11.2;q11.2) translocation found in human synovial sarcoma. Nat. Genet. 7:502-508(1994).	
RT	[2]	
RL	SEQUENCE OF 374-379 FROM N.A. (SSXT-SX2 FUSION PROTEIN).	
RX	MEDLINE-96094743; PubMed-7495284;	
RA	Filgman I., Lonardo F., Jhanwar S.C., Gerald W.L., Woodruff J., Ladanyi M.; "Molecular diagnosis of synovial sarcoma and characterization of a variant SYT-SX2 fusion transcript." Am. J. Pathol. 147:1592-1599(1995).	
CC	-1- TISSUE SPECIFICITY: FAIRLY UNIQUITUOUSLY EXPRESSED. EXPRESSED IN SYNOVIAL SARCOMAS AND IN OTHER HUMAN CELL LINES. THE FUSION GENES SSXT-SX1 AND SSXT-SX2 ARE EXPRESSED ONLY IN SYNOVIAL SARCOMAS.	
CC	-1- DISEASE: PARTICIPATES IN A T(X;18)(p11.2;q11.2) CHROMOSOMAL TRANSLOCATION THAT IS SPECIFICALLY FOUND IN MORE THAN 80% OF SYNOVIAL SARCOMA AND PRODUCES THE SSXT-SX1 OR SSXT-SX2 FUSION PRODUCTS. THESE HYBRID PROTEINS ARE PROBABLY RESPONSIBLE FOR TRANSFORMING ACTIVITY. HETEROGENEITY IN THE POSITION OF THE BREAKPOINT CAN OCCUR (LOW FREQUENCY).	
CC	-1- SIMILARITY: STRONG, TO MOUSE SSXT.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC	EMBL: XJ9201; CAA55792.1; ALT_INIT.	
CC	DR EMBL: ST9894; AAB5674.1; ALT_TERM.	
CC	DR MM: 600192; -.	
CC	DR Chromosomal translocation: Proto-oncogene; Repeat.	
FT	SITE 335 336 BREAKPOINT FOR TRANSLOCATION TO FORM THE SSXT-SX1 FUSION PROTEIN (RARE).	
FT	SITE 379 380 BREAKPOINT FOR TRANSLOCATION TO FORM THE SSXT-SX1 FUSION PROTEIN (RARE).	
FT	SITE 50 53 SH2-BINDING (POTENTIAL).	
FT	SITE 343 346 SH2-BINDING (POTENTIAL).	
FT	DOMAIN 361 370 SH3-BINDING (POTENTIAL).	
FT	SITE 382 385 SH2-BINDING (POTENTIAL).	
FT	DOMAIN 313 338 2 X 13 AA IMPERFECT TANDEM REPEATS.	
FT	REPEAT 313 325 1.	
FT	REPEAT 326 338 2.	
FT	REPEAT 95 99 POLY-PRO.	
FT	DOMAIN 317 320 POLY-GLN.	
SQ	SEQUENCE 387 AA; 42217 MW; ALD64C2C588C532A CRC64;	
DR	EMBL: X93357; CAA63733.1; -.	Query Match 19.8%; Score 139; DB 1; Length 418;
DR	MGI: 107708; Ssxt.	Best Local Similarity 36.9%; Pred. No. 0.00011;
KW	Repeat.	
FT	SITE 374 377 SH2-BINDING (POTENTIAL).	
FT	DOMAIN 392 401 SH3-BINDING (POTENTIAL).	
FT	SITE 413 416 SH2-BINDING (POTENTIAL).	
FT	DOMAIN 344 369 2 X 13 AA IMPERFECT TANDEM REPEATS.	
FT	REPEAT 344 356 1.	
FT	REPEAT 357 369 2.	
FT	DOMAIN 95 99 POLY-PRO.	
FT	348 351 POLY-GLN.	
SQ	SEQUENCE 418 AA; 45891 MW; D32B66AD9396CA2 CRC64;	

Db	147	-QSYIPPPGQYQQNQYNSSGGGGGGGGNYQDQSSMSGGGGGGYGNQDQSSGGGS	205	RA	Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwynn M.,
Qy	115	AGFQFQSQG 123		RA	Doughterty B., Tomb J.-F., Fleischmann R.D., Richardson D., Richardson D.,
	:	:		RA	Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
Db	206	GGYGQDRG 214		RA	van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
				RA	Utterback T., Wathey L., McDonald L., Artlach P., Bowman C.,
				RA	Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
				RA	Smith H.O., Ventier J.C.;
RESULT	13	YBM6_YEAST	STANDARD; PRT; 128 AA.	RT	"Genomic sequence of a Lyme disease spirochaete, <i>Borrelia burgdorferi</i> ."
ID	YBM6 YEAST	STANDARD;	PRT;	128 AA.	RT
AC	P38216				Nature 390:580-586(1997).
DT	01-OCT-1994 (Rel. 30, Created)				-I - FUNCTION: IF-2, ONE OF THE ESSENTIAL COMPONENTS FOR THE INITIATION
DT	01-OCT-1994 (Rel. 30, Last sequence update)				OF PROTEIN SYNTHESIS IN VITRO, PROTECTS FORMYLMEHTIONYL-tRNA FROM
DE	01-NOV-1997 (Rel. 35, Last annotation update)				SPOONTANEOUS HYDROLYSIS AND PROMOTES ITS BINDING TO THE 30S
DE	HYPOTHETICAL 14.6 kDa PROTEIN IN TRP1-KAP104 INTERGENIC REGION.				RIBOSOMAL SUBUNITS. IT IS ALSO INVOLVED IN THE HYDROLYSIS OF GTP
GN	YBR016W OR YBR022Z				DURING THE FORMATION OF THE 70S RIBOSOMAL COMPLEX (BY SIMILARITY).
OS	Saccharomyces cerevisiae (Baker's yeast).				-I - SUBCELLULAR LOCATION: CYTOPLASMIC
OC	Eukaryota; Fungi; Ascomycota; Saccharomycetidae; Saccharomycetes;				-I - SIMILARITY: BELONGS TO THE IF-2 FAMILY.
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes;				CC
OX	NCBI_TAXID=4932;				CC
RN	[1]				CC
RP	SEQUENCE FROM N.A.				CC
RC	STRAIN=S28C;				CC
RA	Entian K.-D., Koetter P., Rose M., Li Z., Thermann R., Brendel M.,				CC
RA	Baur A., Boles E., Miosga T., Schaaff-Gerstenschlaeger I.,				CC
RA	Zimmermann F.K.,				CC
RL	Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.				CC
CC					CC
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				CC
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				CC
CC	the European Bioinformatics Institute. There are no restrictions on its				CC
CC	use by non-profit institutions as long as its content is in no way				CC
CC	modified and this statement is not removed. Usage by and for commercial				CC
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).				CC
CC					CC
CC					CC
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				CC
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				CC
CC	the European Bioinformatics Institute. There are no restrictions on its				CC
CC	use by non-profit institutions as long as its content is in no way				CC
CC	modified and this statement is not removed. Usage by and for commercial				CC
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).				CC
DR	EMBL; A0084558.1; -.				CC
DR	EMBL; A0084558.1; -.				CC
DR	TIGR; BB0801; -.				CC
DR	InterPro; IPR000795; GTP_EFTU.				CC
DR	InterPro; IPR000178; IF2.				CC
DR	ProDom; PD186100; IF2; 1.				CC
DR	PROSITE; PS01176; IF2; FALSE_NEG.				CC
KW	Initiation factor; Protein biosynthesis; GTP-binding;				CC
KW	Complete proteome.				CC
FT	DOMAIN 383 531				CC
FT	NP_BIND 389 396				CC
FT	GTP (BY SIMILARITY).				CC
FT	NP_BIND 439 439				CC
FT	GTP (BY SIMILARITY).				CC
SQ	SEQUENCE 489 492				CC
SQ	SEQUENCE 882 AA; 97795 MN; 0FA2D065B543C41 CRC64;				CC
Query	Match 19.3%; Score 135; DB 1; Length 128;				CC
Best	Local Similarity 40.7%; Pred. No. 7.6e-05;				CC
Matches	44; Conservative 4; Mismatches 38; Indels 22; Gaps 6;				CC
Qy	7 GNNQNYQQS-----QNNQOQNQNYQQYQAYNAQAOQAGGYQNYQGYS-- 53				CC
Db	8 GGTAGEKSQYSPSPNPPPSAHQNKTQERGYPQQQQYQQQQQHPG-YYNQOGYNQ 65				CC
Qy	54 GYQOGYQQYQNPDAQYQQ-YNPGQYQQYNPQGGQYQQNPQGRGN 100				CC
Db	66 GYNQOGYNQ---QG3NQGQYNNQDQ3NQGQHQPQYVQOPPPQ-RGN 107				CC
RESULT	14				CC
ID	IF2_BORBU	STANDARD;	PRT;	882 AA.	CC
AC	051741				CC
DT	15-DEC-1998 (Rel. 37, Created)				CC
DT	15-DEC-1998 (Rel. 37, Last sequence update)				CC
DT	20-AUG-2001 (Rel. 40, Last annotation update)				CC
DE	TRANSLATION INITIATION FACTOR IF-2.				CC
GN	INF_B OR B0801				CC
OS	Borrelia burgdorferi (Lyme disease spirochete).				CC
OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.				CC
OC	NCBI_TAXID=139;				CC
RN	[1]				CC
RP	SEQUENCE FROM N.A.				CC
RC	STRAIN=ATCC 35210 / B31;				CC
RX	MEDLINE-#90065943; PubMed-#9403685;				CC
RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,				CC
RA	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				CC
RESULT	15				CC
RD_RAT	RD_RAT STANDARD; PRT; 353 AA.				CC
ID	RD_RAT				CC
AC	O9JJ54; O9JJ53; O9JJ52; O9JJ51;				CC
DT	20-AUG-2001 (Rel. 40, Created)				CC
DT	20-AUG-2001 (Rel. 40, Last sequence update)				CC
DT	20-AUG-2001 (Rel. 40, Last annotation update)				CC
DE	HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN D0 (AU-RICH ELEMENT RNA-BINDING PROTEIN 1).				CC
GN	HNRNP OR AU1P.				CC
OS	Rattus norvegicus (Rat).				CC
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				CC

OX_NCBI_TAXID=10116;
 RN_11];
 RPSEQUENCE FROM N.A.
 RC
 RA TISSUE=Kidney;
 RA Arao Y., Kikuchi A.;
 RA "Differential expression of AUPL isoforms in rat tissues.";
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RL
 CC -1- FUNCTION: Binds with high affinity to RNA molecules that contain
 AU-RICH ELEMENTS (ARES) found within the 3'-PRIME UNTRANSLATED
 REGIONS OF MANY PROTOONCOGENES AND CYTOKINE MRNAs (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; COMPONENT OF RIBONUCLEOSOMES (BY
 SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; 1/P45 (SHOWN HERE), 2/P42, 3/P40
 CC AND 4/P37; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.ebi-sib.ch/announce/>
 or send an email to license@ebi-sib.ch).
 CC
 DR EMBL: AB046615; BAB03465; 1;
 DR EMBL: AB046616; BAB03466; 1;
 DR EMBL: AB046617; BAB03467; 1;
 DR EMBL: AB046618; BAB03468; 1;
 DR InterPro: IPR00504; RRM.
 DR Pfam: PF00076; rrm; 2.
 DR SMART: SM00360; RRM; 2.
 DR PROSITE: PS00010; RRM; 2.
 DR PROSITE: PS000102; RRM; 2.
 DR KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat;
 KW Alternative splicing.
 FT DOMAIN 11 45 ALA-RICH.
 FT DOMAIN 95 177 RNA-BINDING (RRM) 1.
 FT DOMAIN 180 259 RNA-BINDING (RRM) 2.
 FT DOMAIN 268 345 GLY-RICH.
 FT DOMAIN 292 330 TYR-RICH.
 FT VARSPLIC 77 95 MISSING (IN ISOFORM 2 AND ISOFORM 4).
 FT VARSPLIC 283 332 GPSNONQHGYSWNGQYGSYNSDQYGGYGYGDSYGSYNGVYGGDSN->D (IN ISOFORM 3 AND ISOFORM
 FT 4).
 SQ SEQUENCE 353 AA; 38192 MW; 6190BE1006F07D0B CRC64;

Query Match 18.3%; score 128; DB 1; Length 353;
 Best Local Similarity 38.5%; Pred. No. 0.00072; Gaps 1;
 Matches 30; Conservative 5; Mismatches 33; Indels 10; Gaps 1;
 QY 17 SONGNOOQGGINRRIQYQAYYQAQPAQGYYIQNYQGSGYQOGGYYQQNPDAQYQQQNPD 76
 QY 285 SQQNWNQGYSWNGQY-----GSYGSNSQGYYGGYDYTGNSVYGGDVSNQ 334
 QY 77 GGYQQYNPQGQYQQNP 94
 QY 335 SGYGKVRQRGHONSYK 352